Thesis for a Ph. D. Degree

An IASI channel selection for UM 1D-Var and its impact on precipitation forecast over East Asia

UM 일차원 변분법을 이용한 IASI 채널 선정과 동아시아 지역 강수 예보 정확도에 미치는 영향

August 2017

School of Earth and Environmental Sciences

Graduate School

Seoul National University

Young-Chan Noh

An IASI channel selection for UM 1D-Var and its impact on precipitation forecast over East Asia

By
Young-Chan Noh

A Dissertation submitted to the Faculty of the
Graduate School of the Seoul National University
in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

Degree Awarded:
August 2017

Advisory committee:

Professor Gyu-Ho Lim, Chair
Professor Byung-Ju Sohn, Advisor
Professor Seok-Woo Son
Doctor Sangwon Joo
Doctor Yoonjae Kim

이학박사학위논문

UM 일차원 변분법을 이용한 IASI 채널 선정과 동아시아 지역 강수 예보 정확도에 미치는 영향

An IASI channel selection for UM 1D-Var and its impact on precipitation forecast over East Asia

2017 년 8 월

서울대학교 대학원
지구환경과학부
노 영 찬

Abstract

Firstly we attempted to understand the analysis performance and associated error characteristics of Unified Model (UM) analysis, aimed at improving initial conditions for better data assimilation system. In doing so, comparison was made between UM analyses at the Met Office and the Korea Meteorological Administration (KMA) against collocated Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN) radiosonde observations for January–December 2012. The European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data were also compared. Although temperature fields in the three analyses are in good agreement with the radiosonde observations, humidity forecasts appear to be less reliable, suggesting that improvement of humidity field is needed in the UM data assimilation system..

We presumed that the UM analysis and its forecast can be improved by running Infrared Atmospheric Sounding Interferometer (IASI) data assimilation with IASI channels different from current operationally used 183 channels. We do this by re-selecting new channels from operationally available 314 channel of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). In selecting channels, we calculated the impact of the individually added channel on the improvement in the analysis outputs from a one-dimensional variational analysis (1D-Var) for the UM data assimilation system at the Met

Office, using a newly defined channel score index (CSI) as a figure of merit. And then 200 channels were orderly selected by counting individual channel's CSI contribution. Compared with the operationally used 183 channels for the UM at the Met Office, the new set shares 149 channels while other 51 channels are new.

The impact of newly selected 200 channels was assessed via experimental trial runs using the UM global assimilation system. The new channels had an overall neutral impact in terms of improvement in forecasts, as compared with results from the operational channels. However, upper-tropospheric moist biases shown in the control run with operational channels were significantly reduced in the experimental trial with the newly selected channels. The reduction of moist biases was mainly due to the additional water vapor channels, which are sensitive to the upper-tropospheric water vapor. In addition, experiments have been conducted for assessing how new channels might improve the precipitation forecast over East Asia where distinct warm type heavy rainfall prevails. Experiments indicate that warm-type heavy rainfall appears to be better forecasted with a new set of IASI channels.

Keyword: Hyper-spectral infrared sounder, channel selection, data assimilation, precipitation, East Asia

Student Number: 2009-20367

Table of Contents

List of Tables

List of Figures

1. Introduction

Numerical weather prediction (NWP) models are extensively used for predicting the future state of the atmosphere, given initial conditions determined from observations. A range of atmospheric phenomena, such as flooding events and tropical cyclones which result in large impacts on society, are possible to predict (Warner, 2011). Particularly, in recent years, the improved weather forecasting skills of NWP models have been partly due to a realistic representation of the atmosphere at the initial state of forecasting. In fact, accurate initial conditions are one of the most important factors for improving the prediction, since the time integration in NWP models fundamentally is an initial-value problem (Lorenc, 1986; Kalnay, 2003; Brown et al., 2012). Hence, many operational weather forecasting centers focus on improving the initial atmospheric state using improved data assimilation schemes and better/more observations of various atmospheric parameters (Evensen, 1994; Lorenc et al., 2000; Andersson et al., 2007; Rawlins et al., 2007; Clayton et al., 2013).

With advent of satellite measurements and development of data assimilation techniques, the impact of satellite observations in numerical weather forecasts has been substantially increased for last three decades. Bauer et al. (2015) showed that 3-day European Centre for Medium-range Weather Forecasts (ECMWF) forecast skill based on the 500-hPa height anomaly correlation in the Northern Hemisphere

has evolved from mid-80 to 98% for last three decades. The forecasting skill in the Southern Hemisphere, which was much lower than in the Northern Hemisphere three decades ago, is now as good as in the Northern Hemisphere. This reduced gap in the forecast skill between the hemispheres manifests the importance of better satellite observations in spectrum, space and time, and the assimilation scheme to optimally combine observations and forecasts. With impressive improvement of satellite observations and assimilation schemes, satellite-measured radiances became the most important observations for the weather forecasting (English et al., 2000; Kelly and Thépaut, 2007; Dee et al., 2011; Lorenc and Marriott, 2014).

Among many different types of satellite data, high-spectral infrared sounder measurements are particularly interesting because high spectral measurements contain rich information on the vertical structure of atmospheric temperature and water vapor. In reality, these sounder measurements give largest impact on the weather forecasting at the UKMO (Hilton et al., 2009; Joo et al., 2013). In this study, we focus on measurements from the Infrared Atmospheric Sounding Interferometer (IASI) instrument onboard the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) MetOp satellites. The IASI instrument senses the emitted radiances at 8461 channels over the infrared spectral range of 645–2760 cm^{-1} (15.5–3.6 μm) at a resolution of 0.5 cm^{-1},

providing information about atmospheric temperature, humidity, cloud variables, and other trace gases. The instrument field of view (FOV) comprises a 2 × 2 array of circular pixels of 0.8 degree angular diameter (about 12 km at nadir) (Simèoni et al., 1997).

It may be desirable to use all of 8461 IASI channels in the data assimilation if measurement errors are within the user-required preset values. However, the direct radiance assimilation of all IASI channels may be practically impossible due to the high data volume and heavy computational burden (Rabier et al., 2002; Collard, 2007). Furthermore, the current data transmission system operated by EUMETSAT may not disseminate all IASI channel data due to the data transmission costs. Instead only a small subset of the whole channel data is distributed by the EUMETSAT via the Global Telecommunication System (GTS) for the operational use at the NWP centers.

Thus, the channel selection is much desired because the small subset should contain the meteorological information as much as in the whole 8461 channels while substantially reducing the computational burden and data transmission cost. Total 300 channels were selected by Collard (2007). Based on such selection, 314-EUMETSAT-channel data have been operationally transmitted, including data at the Centre National d'Etudes Spatiales (CNES) 14 monitoring channels. In the operational Unified Model (UM) system, 183 channels among 314 EUMETSAT

channels have been used in the observation processing system (OPS) for quality-control. But only 138 channels have been assimilated into the NWP model because of technical problems associated with the UM system at the time when IASI data were first introduced in 2007 (Weston, 2011). With an improved model and advanced radiative transfer modeling in the assimilation system after this first use, the technical problems met in the first attempt to select IASI channels may be alleviated. Nevertheless, there have been no attempts to add different channels or evaluate each channel's contribution in the current advanced assimilation system of the UM. Thus, in this study, we attempt to select new channels from 314 EUMETSAT channels, which can replace the current 183 channels for better assimilation results.

Rabier et al. (2002) reviewed various channel selection methods for the high-spectral infrared sounder and concluded that Rodgers (2000) is the most optimal method. In fact, Collard (2007) used Rodgers (2000) method for his 300 channel selection. In Collard (2007), selection was made at each important spectral bands like CO_2, O_3, and H_2O absorption bands. It is because the degree of freedom for signal (DFS) and the entropy reduction (ER) in Rodgers (2000) are heavily influenced by water vapor channels (especially upper tropospheric water vapor channels). But, because of no objective way to evaluate a relative merit of each band, total 300 number of channels were subjectively/evenly chosen by adding

them all, after selecting channels at each absorption band. Thus, it is of interest to develop an objective way to evaluate each channel's merit across the entire infrared spectrum.

In this study, the research contents were divided into two parts. First, in order to understand the error characteristics of initial condition (i.e., analysis) prior to selecting the new IASI channels, the initial conditions from the UM data assimilation at the Met Office and the Korea Meteorological Administration (KMA) were evaluated against collocated radiosonde observations for January–December 2012. Furthermore, the ECMWF reanalysis data as “proxy” true data in weather forecasting were also validated. The radiosonde observations as a reference were obtained from the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN).

And then, we devise a way to select channels with an aid of a one-dimensional variational analysis (1D-Var), which can be used as a means to select channels more objectively. Main aims are reducing problems caused by such subjective or empirical decision on the channel selection. In addition, the Rodger's selection method adopted by Collard (2007) adopts fixed Jacobians while the 1D-Var method devised in this study employs Jacobians variable with 1D-Var iterations. Thus, it is also interesting to examine how Rodger's method behaves if variable Jacobians (resulted from the use of 1D-Var approach) are applied. In so doing,

results from Rodger's approach with variable Jacobians will be compared with results from the proposed 1D-Var method.

In order to examine how the new IASI channels selected in this study improve the global forecasting accuracy, in comparison to the current use of 183 channels in the operational UM data assimilation, we will conduct trial experiments with two different sets of IASI channels. In these experiments, however, different from an observing system simulation experiment (OSSE) studies, all the same observations will be kept as inputs to the system except IASI observations.

In addition to the impact on the global forecasting accuracy, it is also of interest to evaluate the impact of new IASI channels on the heavy precipitation forecasts over East Asia. It has been noted that heavy rainfall events over the Korean peninsula and surroundings are significantly different from those occurred over the Great Plains of the United States (Sohn et al., 2013; Song and Sohn, 2015). While, in general, the strong rainfall occurs under convectively unstable conditions linked to the deep convective cloud system (referred to as cold-type rain), frequent heavy rainfalls over East Asia are associated with relatively low-level cloud which may contain a small amount of ice crystal (referred to as warm-type rain). Due to the unique cloud-rain system over East Asia, the warm-type rain is not well described in the NWP model (Hong, 2004). In this study we are interested in examining how the warm or cold type of rainfall events are forecasted

with the use of new IASI channel selection results. By conducting trial experiments with two different sets of IASI channels, similar to the trial experiment for global forecast accuracy, we will assess the impact of new IASI channels on the forecast accuracy of heavy rainfall events over East Asia. The results obtained in this study will certainly enhance our understanding of how the channel selection can extract most out of IASI observations and thus improve our level of forecasting capability.

2. Used data and numerical weather prediction (NWP) model

2.1. IASI infrared hyperspectral instrument

The Infrared Atmospheric Sounding Interferometer (IASI) is an infrared Fourier-transform spectrometer and is a key payload element of the MetOp series of European meteorological polar-orbit satellites (Simèoni et al., 1997). IASI represents an advanced atmospheric sounding system, along with its partner microwave instrument Advanced Microwave Sounding Unit (AMSU-A). The first flight was launched in 2006 onboard the first European meteorological polar-orbiting satellite, MetOp-A. The second IASI instrument, mounted on the MetOp-B satellite, was launched in September 2012. The third instrument will be mounted

on the MetOp-C satellite with launch scheduled for October–November 2019. The main goal of the IASI instrument is to provide atmospheric emission spectra to derive temperature and water vapor profiles with a vertical resolution of one kilometer and average accuracy of 1 K in temperature and 10% in humidity. In addition, this instrument is used to determine trace gases including ozone, nitrous oxide, carbon dioxide, and methane, as well as land and sea surface temperature and emissivity and cloud properties.

IASI instrument has 8461 channels that are aligned in three bands between 645 cm^{-1} and 2760 cm^{-1} (15.5 μm and 3.63 μm), with a spectral resolution of 0.5 cm^{-1} after apodisation. The spectral sampling interval is 0.25 cm^{-1}. The radiometric calibration performed during the on-board processing provides three real spectra corresponding to each of the three bands of IASI instrument (Table 1). IASI is an across track scanning system with scan range of ±48°20´, symmetrically with respect to the nadir direction. A nominal scan line covers 30 scan positions towards the Earth and two calibration views. One calibration view is into deep space, the other is observing the internal black body. The scan starts on the left side with respect to the flight direction of the spacecraft (Fig. 1). The effective field of view (EFOV) is the useful field of view at each scan position. Each EFOV consists of a 2 x 2 matrix of so-called instantaneous fields of view (IFOV). Each IFOV has a diameter of 14.65 mrad which corresponds to a ground resolution of 12 km at nadir

and a satellite altitude of 819 km. The 2 x 2 matrix is centered on the viewing direction. The instrument points spread function (PSF) is defined as the horizontal sensitivity within an IFOV. The radiometric noise performance is specified in terms of Noise Equivalent Temperature difference (NEdT) at a blackbody temperature of 280 K. For difference scene temperatures, the specification corresponds to the same Noise Equivalent Brightness difference (NEdL). It includes all the noise sources originating from both the electronic and the optics. The required values are 0.25 for band 1 (645–1210 cm^{-1}), 0.25 for band 2 (1250–1650 cm^{-1}), and 0.4 for band 3 (2150–2450 cm^{-1}) (Simèoni et al., 1997; EUMETSAT, 2012).

The spectral range of IASI measurement includes gaseous absorption bands of water vapor in the 6.7 μm, ozone in the 9.6 μm, and carbon dioxide in the 4.3 μm and 15 μm. In addition, the range also covers an infrared window region where the atmospheric gases are nearly transparent in the vicinity of 11 μm and surface information could be obtained without the contamination by atmosphere. From the atmospheric information on gaseous absorption and window regions over IASI spectral range, the various atmospheric variables, such as temperature and water vapor profiles, have been produced and distributed with the purpose of improvement of weather forecasting and climate researches.

Table 1. Three spectral bands of IASI instrument.

Band	**Wavenumbers (cm^{-1})**	**Wavelength (μm)**
1	645.0 – 1210.0	8.26 – 15.50
2	1210.0 – 2000.0	5.00 – 8.26
3	2000.0 – 2760.0	3.62 – 5.00

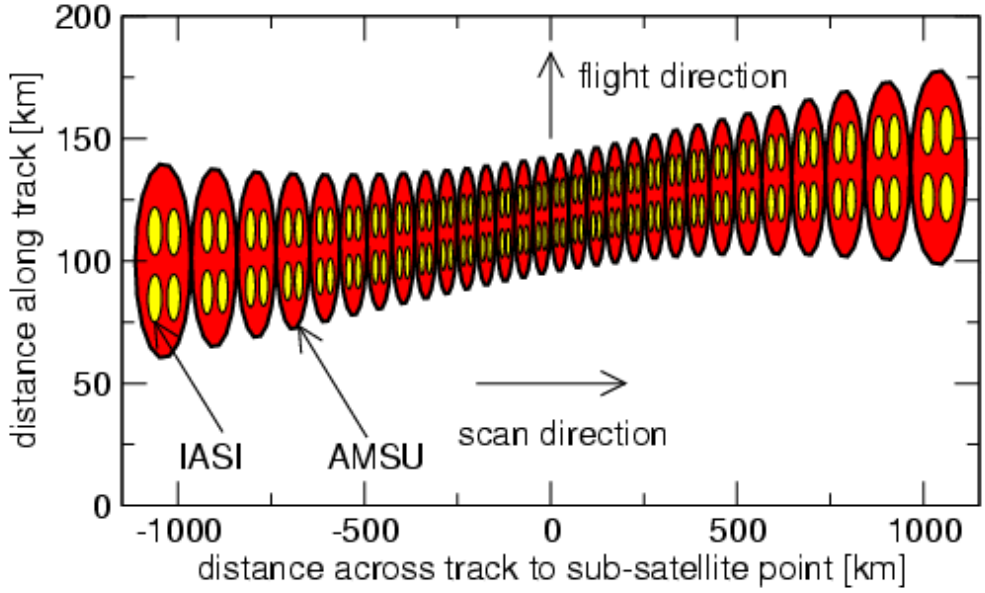

Figure 1. IASI scan geometry and the collocation between IASI (yellow) and AMSU (red) instrument.

2.2. The Unified Model

In May 2010, the Korea Meteorological Administration (KMA) introduced the UM-based global and regional NWP system and its data assimilation system from the United Kingdom Met Office to improve the NWP performance and provide the high-quality weather service to the public (Lee et al., 2012). The Unified Model (UM) is the name given to the suite of atmospheric and oceanic numerical modelling software developed and used at the Met Office. The formulation for the NWP model supports global and regional domains and is applicable to a wide range of temporal and spatial scales that allow it to be used for both numerical weather forecasting and climate modelling as well as a variety of related research activities. A typical run of the UM system consists of an optional period of data assimilation followed by a prediction phase. Forecasts from a few hours to a few days ahead are required for numerical weather prediction while for climate modelling the prediction phase may be for tens, hundreds or thousands of years. The choice of horizontal and vertical resolution of the UM system can be varied by the user, but the resolution is constrained by the available computing power and a number of standard resolutions tend to be used.

Basically, the UM comprises a non-hydrostatic, hybrid height formulation with semi-Lagrangian advection and finite differences in a regular latitude-longitude grid (Davies et al., 2005; Rawlins et al., 2007). For atmosphere

prediction, the UM system uses a set of equations that represent the time-evolution of the atmosphere. The used equations are solved for the motion of a fluid on a rotating, almost-spherical planet. The main atmospheric variables are the zonal and meridional components of the horizontal wind, potential temperature, and specific humidity. To solve the system of equations, a number of approximations and assumptions have to be made. In particular, the UM's atmospheric prediction scheme is based on the so-called hydrostatic primitive equation. Most other NWP models make use of shallow-atmospheric approximation that derives its name from the fact that the depth of the atmosphere is much smaller than the radius of the Earth. However, The UM does not use this approximation, so this makes the equations a little more complicated to solve by adding in extra terms which can be important when planetary scale motions are considered (Met Office, 1998).

In order to predict the future state of the atmosphere based on current atmosphere condition from the UM data assimilation system, atmospheric and oceanic data assimilation is performed to produce the analysis data as the initial condition for weather forecasting. In the data assimilation process, a number of available observation data including surface-based observation and various satellite measurements are utilized to more accurately track developing weather systems. The Met Office had operated three-dimensional variational data assimilation (3D-Var) from 1999, when it improved upon an earlier nudging

method as an analysis correction, particularly with the introduction of satellite sounding data from Advanced TIROS operational Vertical Sounder (ATOV) instruments (English et al., 2000). On late 2004, the Met Office global data assimilation was switched from 3D-Var to four-dimensional variational data assimilation (4D-Var). The UM global data assimilation system comprises an update cycle with a data window of 6 hours, such that observations are ingested from 2100–0300, 0300–0900, 0900–1500, 1500–2100 UTC, with a nominal T+0 analysis time in the middle of each period (Rawlins et al., 2007). And the UM global model for 10-day forecast runs at 00 UTC and 12 UTC and the 3-day forecasting is also produced at 06 UTC and 18 UTC. Recently, the Met Office global UM model has implemented the hybrid ensemble/four-dimensional variational (4D-Var) data assimilation system. The hybrid ensemble/4D-Var uses the extended control variable technique to implement a hybrid background error covariance that combines the standard climatological covariance with a covariance derived from the 23 ensemble members produced from MOGREPS-G that is the operational ensemble prediction system (EPS) at the Met Office (Clayton et al., 2013).

3. Evaluation of Unified Model analysis and ERA-Interim reanalysis data

Although the analysis outputs are now routinely used for weather forecasting and diagnostic studies, they require validation by in situ measurements, such as radiosonde data, as doing so will lead to a better understanding of the analysis performance and associated error characteristics. Such efforts are particularly important when new NWP models are introduced or when model systems including physical schemes and/or data assimilation systems are upgraded.

For the validation of the analysis/forecast performance, reference data are required. Radiosonde observations are often used for validating the vertical structure of atmospheric variables. In the past, the use of radiosonde data for validation has been hampered because of sensor uncertainties and poor documentation regarding the instrument changes, since measurement errors are significantly affected by the sensor design, calibration, and data processing. In addition, the uncertainty of humidity observations was especially large for the upper troposphere in dry and cold conditions (Nash et al., 2011; Moradi et al., 2013). However, as more accurate radiosonde sensors become available, such as in the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN) (Dirksen et al., 2014), the atmospheric profiles of model analysis can be assessed more confidently.

As radiosondes are one of the most important observations in data assimilation systems, some radiosonde measurements from GRUAN stations (prior to final GRUAN processing) are already used in assimilation at operational forecast centers, which implies that the analysis is not fully independent of the GRUAN data. However, considering that the radiosonde measurements are differently assimilated at each forecast center, their impact on the analysis may be different from model to model, although the same raw data are given. Furthermore, the analysis is produced by assimilating various other observations besides radiosonde data. Consequently, the main aim of this study is to examine the overall accuracy of the analysis produced through the data assimilation against GRUAN-measured temperature and humidity profiles as references.

In May 2010, the KMA introduced the UM and associated data assimilation systems developed by the UKMO (Lee et al., 2012). However, to our knowledge, since the KMA UM has been operationally used, no validation regarding its accuracy was published. Thus, it is informative to assess the accuracy of KMA UM analysis compared with UKMO UM analysis, since the UKMO and KMA use the same NWP system but different versions of the model/assimilation system. Furthermore, the ECMWF reanalysis data, which have been commonly used as "proxy" true values in weather forecasting and climate science, are also validated using the same GRUAN dataset.

3.1. Data used and processing

Two UM analysis datasets (from the KMA and UKMO), ECMWF ERA-Interim reanalysis data (ERA-I), and GRUAN radiosonde observations from January–December 2012 are compared. The detailed data information and comparison procedures are given below.

3.1.1. Construction of collocation data

The comparison data were produced by the data assimilation system of the UM-based global model (UM model). Model outputs were obtained from the KMA and UKMO (UM-KMA and UM-UKMO data, respectively). The spatial resolution of both UM models is N512L70 with 1024 × 769 × 70 grid points, approximately equivalent to 25 km grid spacing at mid-latitudes and 70 vertical levels from the surface to 80 km. The model used for data assimilation has a coarser spatial resolution of N216L70 (about 60 km) (KMA, 2012). The UM-KMA and UM-UKMO used the same assimilation system during the validation period, i.e., a four-dimensional variational assimilation (4D-Var) scheme with 6 h update cycles within the assimilation time window from –3 to +3 h centered at the nominal T + 0 analysis time (i.e., 12 a.m., 6 a.m., 12 p.m. and 6 p.m. UTC) (Rawlins et al., 2007; Clayton et al., 2013). In this study, the T + 0 analysis data

are considered as initial conditions and compared with the reference radiosonde data.

The temperature and relative humidity ERA-I data used in this study use a horizontal grid of 0.25° × 0.25° (about 18 km × 25 km at 45° mid-latitude) for 37 standard pressure levels between 1000 and 1 hPa, and are given four times a day (i.e., 12 a.m., 6 a.m., 12 p.m. and 6 p.m. UTC) (Dee et al., 2011). It should be noted that ERA-I data are based on the use of the August 2006 operational NWP system at ECMWF, which is older than the two UM systems (around 2012). Since the radiosonde standard observation times are 12 a.m. and 12 p.m. UTC, the reanalysis data for the same UTC hours are used for collocation.

As a reference, we use radiosonde data from the GRUAN archives. The GRUAN project was designed to provide high-quality upper atmosphere observations to validate various remote sensing products using the Vaisala RS92 sensors (Immler and Sommer, 2011; Wang et al., 2013; Dirksen et al., 2014). Although products from other GRUAN radiosondes are under development, we only use GRUAN Vaisala RS92 data in this study. The GRUAN applies its own processing scheme (version 2) for the data processing, rather than applying the scheme normally used for the real-time processing such as the Vaisala processing. In addition to the data processing at the operational centers whose qualities are not greatly discernible from the those of GRUAN (Dirksen et al., 2014), the GRUAN

processing carefully derives uncertainties of observations; uncertainties of RS92 radiosonde observations made by the GRUAN are 0.6 K (0.15 K) for daytime (nighttime) temperature measurements at 25 km and maximum of about 3%–5% for relative humidity (Dirksen et al., 2014). The geographical locations of the used GRUAN sites are given in Table 2. Observations made at some GRUAN stations (here LIN, NAU, PAY, SOD, and TAT stations of Table 2) directly go to weather centers through the Global Telecommunication System (GTS) for the real-time data processing for the weather forecasts.

The UM and radiosonde data are mapped for the same 37 pressure levels of the ERA-I data and co-locations are made by selecting the grid point values of the NWP outputs closest to the GRUAN observation sites. Since both the UM model and ERA-I outputs have about 25 km spacing at the equator, typically the maximum distance between the radiosonde site and the co-located model data is 12.5 km, resulting in standard deviation increases of a maximum amplitude of 0.05 K for temperature and 0.4% for relative humidity (Sun et al., 2010). This representativeness uncertainty and the influence of the ERA-I coarser model grid (0.75°) need to be considered along with measurement uncertainty when comparing observed and model values.

Although the GRUAN data are well calibrated and carefully maintained, technical and/or artificial errors may be present. To remove possible erroneous

data, we applied a simple quality control scheme to exclude extreme outliers from the statistical mean. If the temperature or moisture for a given level is above or below ±3 σ from the January–December 2012 mean, that level of data would be removed from each profile. Thus, the total number of samples used for the statistical analysis varies with the level. However, there is little variation, except near the surface and above 250 hPa (Fig. 2a). Slightly decreasing sample numbers are noted for 250–50 hPa, while a rapid decrease is found above 50 hPa. The smaller number of samples near the surface is mainly due to the varying altitudes of the observation sites.

Vertical mean profiles of temperature and relative humidity obtained from GRUAN radiosonde observations and associated uncertainties for the January–December 2012 period are given in Figure 2b. Also given in Figure 2b are uncertainties of temperature and moisture profiles. The level of temperature uncertainty gradually increases with height and approaches 0.2 K at around the 50 hPa level and 0.3 K at around the 10 hPa level. At the same time, the mean relative humidity approaches zero above the 100 hPa level. Because of these points, the statistical comparison is made for all available data between the surface and 100 hPa.

Table 2. Information on the Global Climate Observing System Reference Upper-Air Network (GRUAN) radiosonde sites used in this study. Numbers given with names are WMO station numbers used for the GTS.

Code	Name	Country	Location			International Name
			Latitude	Longitude	Altitude	
CAB	Cabauw	Netherlands	51.97˚	4.92˚	1 m	Cabauw (De Bilt) Experimental Site for Atmospheric Research (CESAR)
LIN	Lindenberg (10393)	Germany	52.21˚	14.12˚	110 m	Lindenberg Meteorological Observatory
MAN	Manus	Papua New Guinea	-2.06˚	147.42˚	6 m	Tropical Western Pacific (TWP) Manus Site
NAU	Nauru (91532)	Nauru	-0.52˚	166.92˚	7 m	Tropical Western Pacific (TWP) Nauru Site
NYA	Ny-Ålesund	Norway	78.92˚	11.92˚	5 m	AWIPEV Research Base
PAY	Payerne (06610)	Switzerland	46.81˚	6.95˚	491 m	MeteoSwiss Aerological station Payerne
SGP	Lamont	USA	36.6˚	-97.49˚	320 m	Southern Great Plains (SGP) Central Facility
SOD	Sodankyla (20386)	Finland	67.37˚	26.63˚	179 m	Finnish Meteorological Institute Arctic Research Centre
TAT	Tateno (47646)	Japan	36.06˚	140.13˚	31 m	Tateno Aerological Observatory

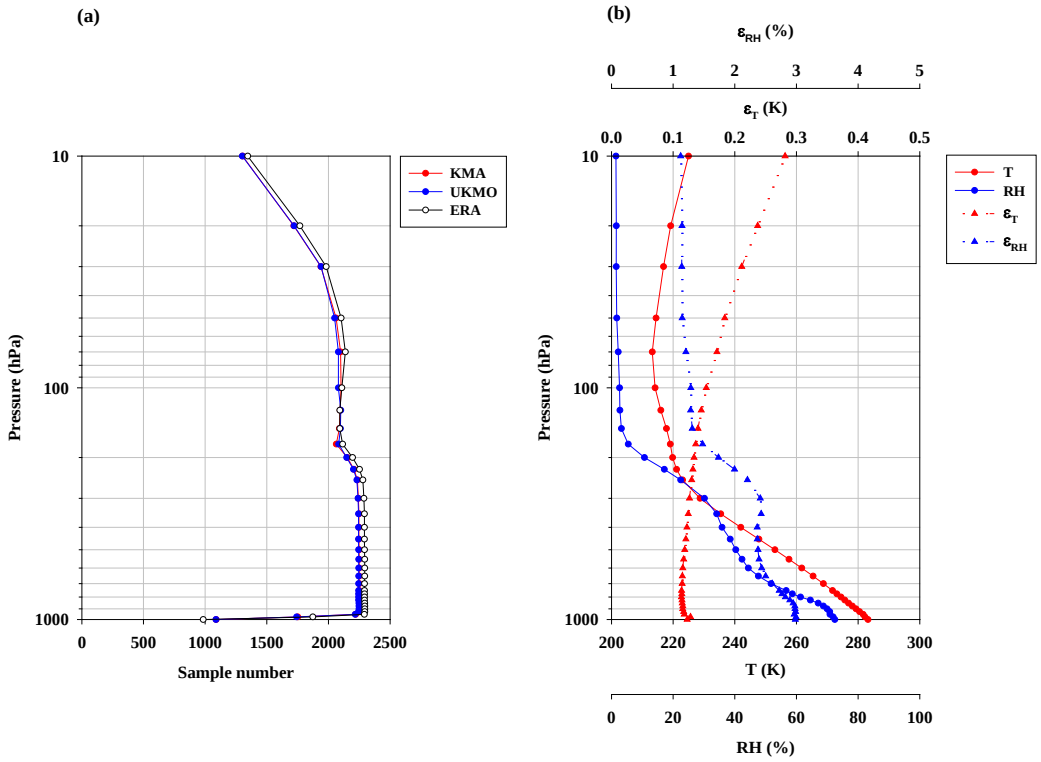

Figure 2. (a) Number of collocated data from the radiosonde observations and numerical weather prediction (NWP) model outputs from the Korea Meteorological Administration (UM-KMA), United Kingdom Met Office (UM-UKMO), and European Centre for Medium-Range Weather Forecasts Interim reanalysis data (ERA-I), and (b) mean temperature (T) and relative humidity(RH) (solid lines with dots) and associated mean uncertainties (ε_T for temperature and ε_{RH} for relative humidity; dashed lines with triangles) profiles from GRUAN observations.

3.1.2. Conversion of the relative humidity

Since the relative humidity is defined as the ratio between the water vapor pressure and the saturation vapor pressure for a given air temperature, the relative humidity varies depending on the definition of saturation vapor pressure (Murphy and Koop, 2005). Considering that super-cooled water droplets coexist with ice particles in the cloud layer above the freezing level, the relative humidity depends on the definition of the saturation vapor pressure, especially for cold environments. The Vaisala RS92 sensor uses a thin-film capacitor as a relative humidity sensor, in which the measured capacitance is proportional to the ambient water vapor concentration in relation to the liquid water (Miloshevich et al., 2006). Therefore, the GRUAN algorithm calculates the saturation vapor pressure in relation to the liquid water using the equation of Hyland and Wexler (1983), even below the freezing temperature. However, in the UM model, it is calculated using a derivative of the Goff-Gratch formula (Goff and Gratch, 1946; Goff, 1957), which estimates the saturation vapor pressure over ice for an ambient air temperature below freezing.

Unlike the GRUAN or the UM, the ECMWF reanalysis uses a modified version of the Tetens formula as a function of temperature (Simmons et al., 1999; Clark and Harwood, 2003). Buck's (1981) coefficients are used for temperatures above 0°C, while for temperatures below −23°C, the coefficients suggested by

Alduchov and Eskridge (1996) are used to calculate the saturation vapor pressure. Between 0°C and −23°C, a quadratic interpolation scheme is used.

To compare the relative humidity under the same framework and prevent discrepancies in the relative humidity due to different definitions of the saturation vapor pressure, we used the same GRUAN's Hyland and Wexler (1983) equation for the saturation vapor pressure-temperature relationship for all four datasets. The profiles of the mean relative humidity for the given temperature profiles are shown in Figure 2b.

3.2. Results and discussion

We first compare the mean statistics of the UM and ERA-I data with the GRUAN radiosonde data, and then compare variations in the NWP model outputs classified by atmospheric moistness. The biases of the NWP model outputs in relation to the radiosonde measurements and their associated root mean square differences (RMSDs) are calculated for each pressure level.

3.2.1. Mean statistics

The mean biases and RMSDs of the two UM (i.e., KMA and UKMO) and ERA-I data against GRUAN radiosonde observations are shown in Figure 3. For

the relative humidity, the normalized biases and associated RMSDs are also plotted, after normalizing the profiles of the error statistics against the radiosonde mean profile given in Figure 2b.

The statistical results for temperature from the three NWP models are similar, with a bias less than 0.5 K in magnitude. Less agreement is found for the mean bias above 200 hPa. The mean biases and RMSDs are nearly identical for KMA and UKMO, with only slight differences above 300 hPa. The RMSDs of ERA-I are slightly larger than for the two UM models (1.0 K vs. 1.3 K at 1000 hPa).

Compared with the temperature, the statistical results for the relative humidity significantly differ between NWP models, especially for the upper troposphere (200–400 hPa), where the ERA-I exhibits the largest wet bias (up to 6%), with a substantial difference between the two UM models. For both UM models and the ECMWF reanalysis data, the RMSD distribution increases with height below the 700 hPa level, but decreases gradually above the 700 hPa level. However, considering the decrease in relative humidity with altitude (Fig. 2), it is also relevant to compare the mean bias and RMSD with the mean profile (i.e., normalized profile). The normalized RMSDs of the relative humidity gradually increase with height from the surface and sharply increase above 300 hPa for the three NWP models. In addition, the normalized biases are close to zero below 400 hPa and increase with height above 350 hPa, where the ERA-I shows the largest

wet bias (up to 0.4). However, results in the upper layers should be less reliable because measurements are more uncertain due to the small amount of water vapor and calibration uncertainty of the RS92 (Dirksen et al., 2014). This increase with height appears to be related with the general tendency toward a lower relative humidity in the upper layers, causing relatively larger errors. The general features of the normalized bias also indicate humid biases in the upper layers for all three NWP models, even though the relative humidity is smaller than in the lower layers.

Unlike the larger mean relative humidity bias in the upper troposphere, the RMSDs tend to be larger in the lower troposphere than for the other two. Interestingly, the RMSDs of the ERA-I are slightly larger than for the other two. This may be partly due to the larger time window of ERA-I for the data assimilation compared to the two models. The ERA-I assimilation system uses 4D-Var as does the UM system, but it includes more observation data, due to the larger assimilation time window of ERA-I compared to UM—longer windows/more data are likely to give slightly worse analysis fits to data. On the other hand, considering that the ERA-I data at 0.25° spatial resolution are interpolated from the original ERA-I data for a reduced Gaussian grid with a horizontal resolution of 0.75°, the slightly worse ERA-I fits may also be caused by the residual mismatch caused by the coarse horizontal resolution of the ERA-I (~79 km) compared to the UM (~25 km). Another reason might be that ERA-I is

a somewhat older (August 2006) version of the ECMWF system.

The slight difference between the UM-KMA and UM-UKMO, in spite of the same NWP model, is also interesting—slightly larger RMSDs for the UM-KMA, especially for the relative humidity, are noted. This difference is likely due to the fact that the UKMO, as the agency that developed the UM, was using an updated version, unlike the KMA (Met Office, 2011; KMA, 2012). Moreover, the mean biases of the relative humidity in the upper troposphere are slightly larger for the UM-UKMO than that for the UM-KMA, even though the UKMO used an updated version including more observation data in the assimilation process. This may be due to the introduction of the new UM system at the KMA between late May and early June 2012. Strong dry biases are found in the old UM-KMA during spring (March–May), while humid biases abruptly appear in summer (June–July) in the new UM-KMA outputs. Thus, we suspect that these clear differences, especially for the humidity profile, are the result of the new version.

To examine this interpretation, we analyze the relative humidity at 200, 250, 300, and 350 hPa, where the differences are substantial. We plot scatter diagrams of the UM-UKMO and UM-KMA relative humidity for 2012 (January–May, Fig. 4) and for the period after the introduction of the new version (June–December, Fig. 5). In January–May 2012, the UM-UKMO is more humid than the UM-KMA. Such biases between two models appear to be removed in the June–December

2012 period (Fig. 3), in which the new version was being used at the KMA. Interestingly, the bundles of the points showing extremely dry conditions (<10% of relative humidity) are hardly seen in the new UM-KMA version. Thus, the differences in the two UM outputs should be attributed to different versions of the global model. The reduced dry biases from the new UM-KMA mainly resulted from the use of a more reasonable humidity control variable and the improvement of the inner loop resolution of the 4D-Var from N144 (90 km) to N216 (60 km) in the assimilation process (KMA, 2012). However, even though a new UM system was introduced at the KMA, the RMSDs for the temperature and relative humidity remained larger than for the UM-UKMO (Fig. 3), as the UKMO used a more improved UM version, even after introducing the new version at the KMA.

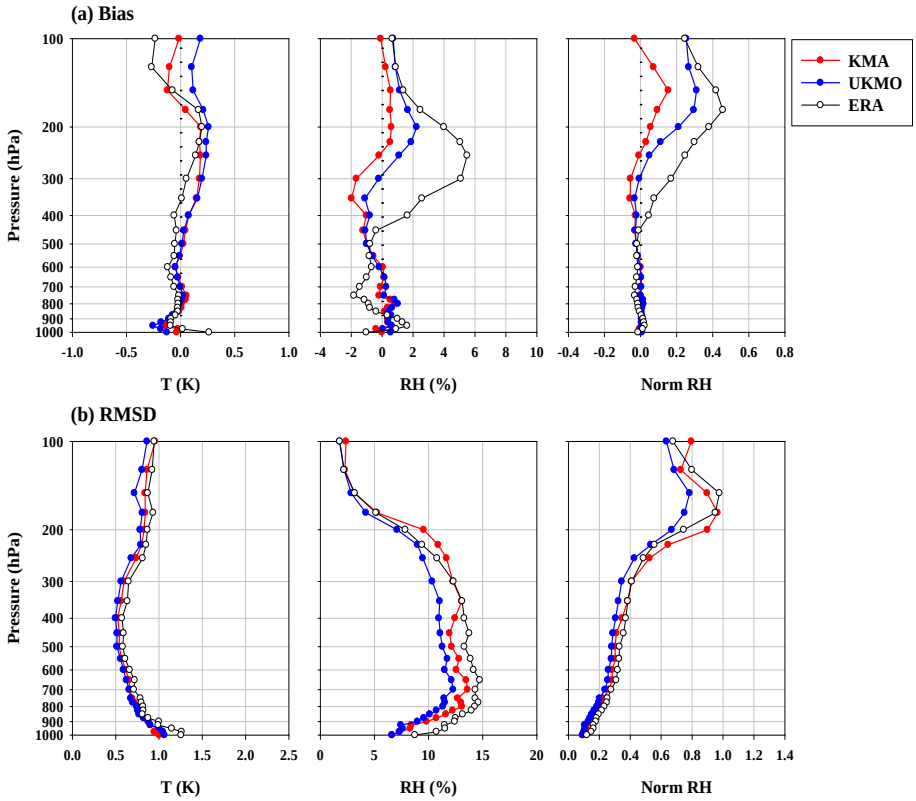

Figure 3. Vertical distributions of the mean biases (top panels) of the temperature, relative humidity, and normalized relative humidity for the Korea Meteorological Administration model output (UM-KMA), United Kingdom Met Office model output (UM-UKMO), and European Centre for Medium-Range Weather Forecasts Interim reanalysis data (ERA-I). Their related root mean square differences (RMSDs) are shown in the bottom three panels.

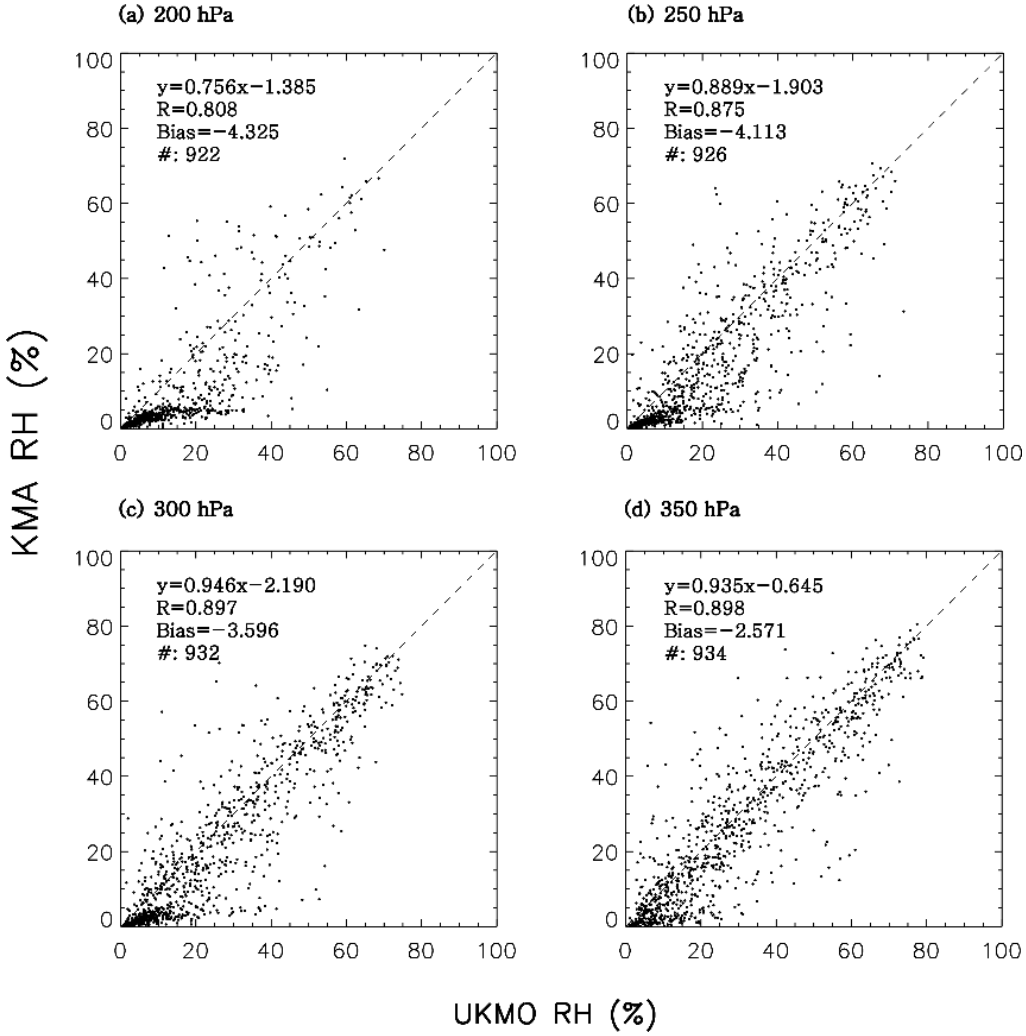

Figure 4. Scatter plots of the relative humidity from the Korea Meteorological Administration model (UM-KMA) versus the United Kingdom Met Office model (UM-UKMO) for (a) 200, (b) 250, (c) 300, and (d) 350 hPa in January–May. The dashed lines represent the perfect matches.

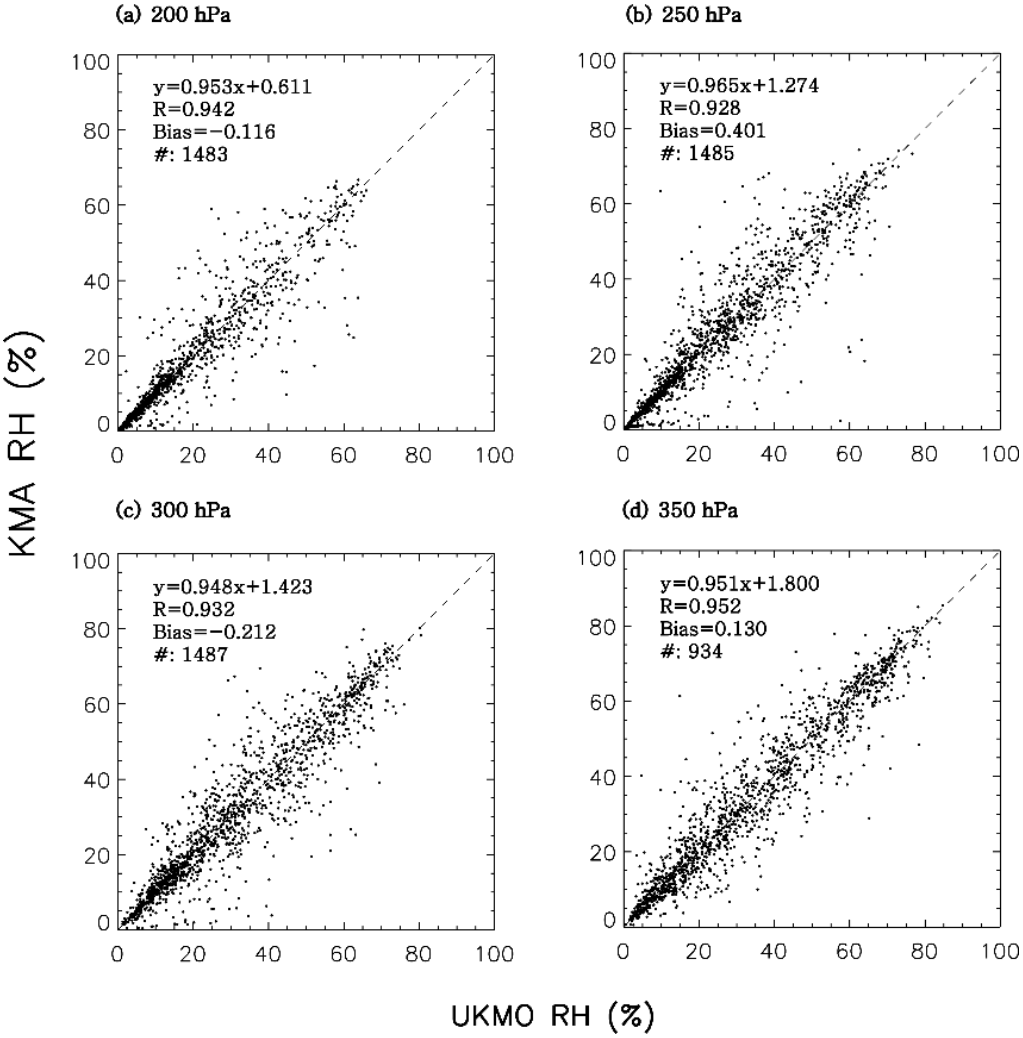

Figure 5. Scatter plots of the relative humidity from the Korea Meteorological Administration model (UM-KMA) versus the United Kingdom Met Office model (UM-UKMO) for (a) 200, (b) 250, (c) 300, and (d) 350 hPa in June–December 2012. The dashed lines represent the perfect matches.

3.2.2. Moistness-depending statistics

Regarding the statistical analysis of the mean bias and RMSDs of the three NWP models compared to the radiosonde observations, larger errors are observed during summer compared to winter (not shown), which might result from the degree of atmospheric moistness. To examine the moistness dependency, the error statistics of the UM analysis and ERA-I reanalysis are examined as a function of the total precipitable water (TPW) between the surface and 100 hPa. Similar analyses were taken for the temperature and moisture retrievals from satellite-based hyperspectral sounder measurements (Kwon et al., 2012; Kwon et al., 2012).

The error statistics for the UM-KMA are given for four GRUAN-based TPW ranges, i.e., <10, 10–20, 20–30 and >30 kg m^{-2} (Fig. 6) — the lowest and highest TPW classes in essence represent polar and tropical regions. The temperature bias tends to increase with increasing moistness, especially for the upper troposphere above 500 hPa. On the other hand, the relative humidity changes from a negative bias for TPW < 30 kg m^{-2} to a positive bias for TPW > 30 kg m^{-2}. Such a bias trend is also found for the normalized relative humidity. The temperature RMSDs show less significant variation with the atmospheric moistness. However, a strong variation with altitude and moistness is evident for the relative humidity RMSDs, which decrease (increase) with an increasing TPW below (above) 700 hPa. The case of TPW > 30 kg m^{-2} shows 15% between 700 and 200 hPa. For the RMSDs

of the normalized relative humidity, the magnitudes tend to decrease with increasing humidity, although such an interpretation is not clear above 400 hPa, suggesting that the moisture fluctuation is larger for a higher humidity in an absolute sense but smaller in a relative sense.

Figure 7 shows the vertical distribution of the relative humidity standard deviation calculated from the radiosonde observations. The standard deviation tends to increase with height until 700 hPa for TPW < 30 kg m^{-2} and until 500 hPa for TPW > 30 kg m^{-2}. Thus, in the lower layers below the 700 hPa level, standard deviations tend to decrease with the increasing TPW. The opposite trend is found above about the 500 hPa level. Those features are similar to the RMSD distribution of the relative humidity (Fig. 6), indicating that the humidity analysis errors are similar to the natural variability.

Figures 8 and 9 show the difference statistics depending on the TPW for the UM-UKMO and ERA-I, respectively. The UM-UKMO results are similar to the UM-KMA (Fig. 6). However, the much larger positive bias in the mean relative humidity of the ERA-I (Fig. 3) is mainly attributed to the humidity biases for TPW < 30 kg m^{-2}, since the biases for TPW > 30 kg m^{-2} are similar to the UM model. Accordingly, clear pattern changes for TPW < 30 kg m^{-2} are visible in the normalized relative humidity bias.

The positive humidity bias of ERA-I around 300 hPa (Figs. 3 and 9) may be

attributed either to the lower saturation vapor pressure from lower temperature or to larger vapor pressure, or to both. Considering that the temperature biases around 300 hPa in the ERA-I are warm (resulting in higher saturation vapor pressure), the moist biases in the upper layer seem to be associated with wetter conditions in the ECMWF analysis in the upper layer. However, caution should be taken because the humidity measurements by radiosonde in the upper tropospheric atmosphere are highly uncertain because of a very small amount of water vapor there, as shown in Figure 2b. Nevertheless, humid bias results are consistent with Hilton et al. (2012), who showed a small wet bias of ECMWF forecasts against IASI brightness temperature measurements.

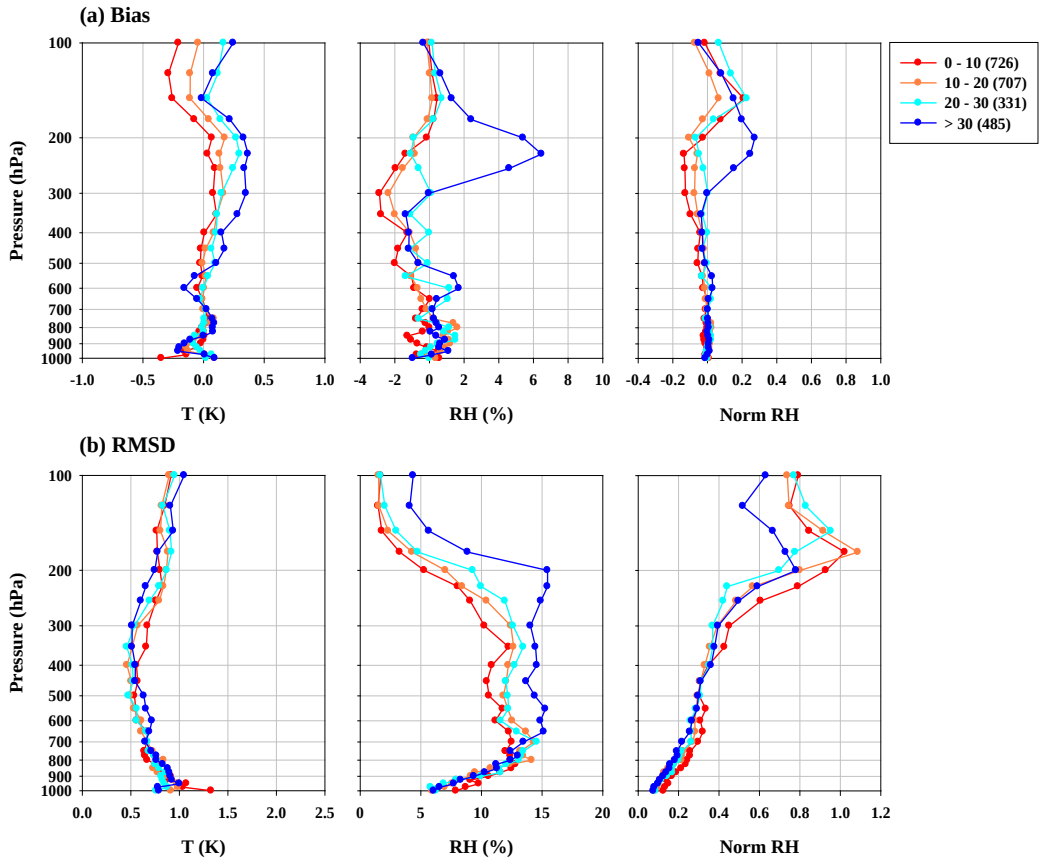

Figure 6. Error statistics of the Korea Meteorological Administration model (UM-KMA) temperature and relative humidity profiles classified according to the total precipitable water between the surface and 100 hPa (TPW, in kg m^{-2}).

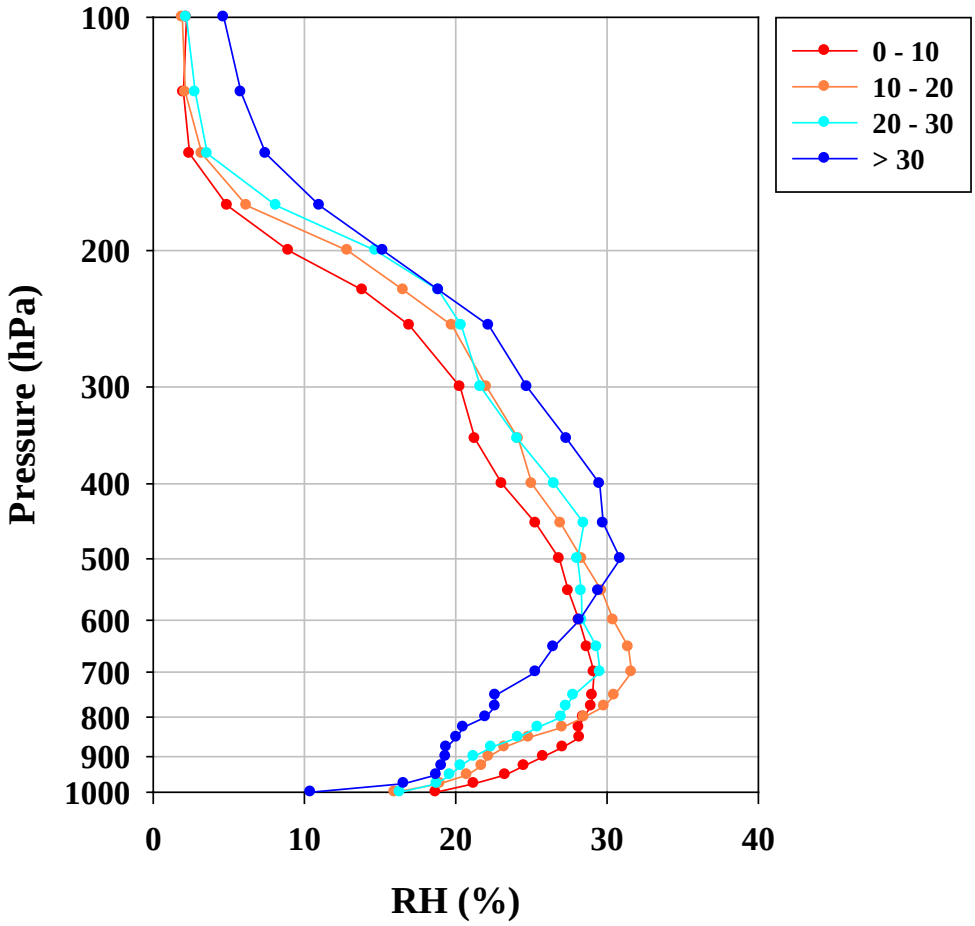

Figure 7. Vertical distribution of the standard deviation of the radiosonde relative humidity profiles classified according to the total precipitable water between the surface and 100 hPa (TPW, in kg m^{-2}).

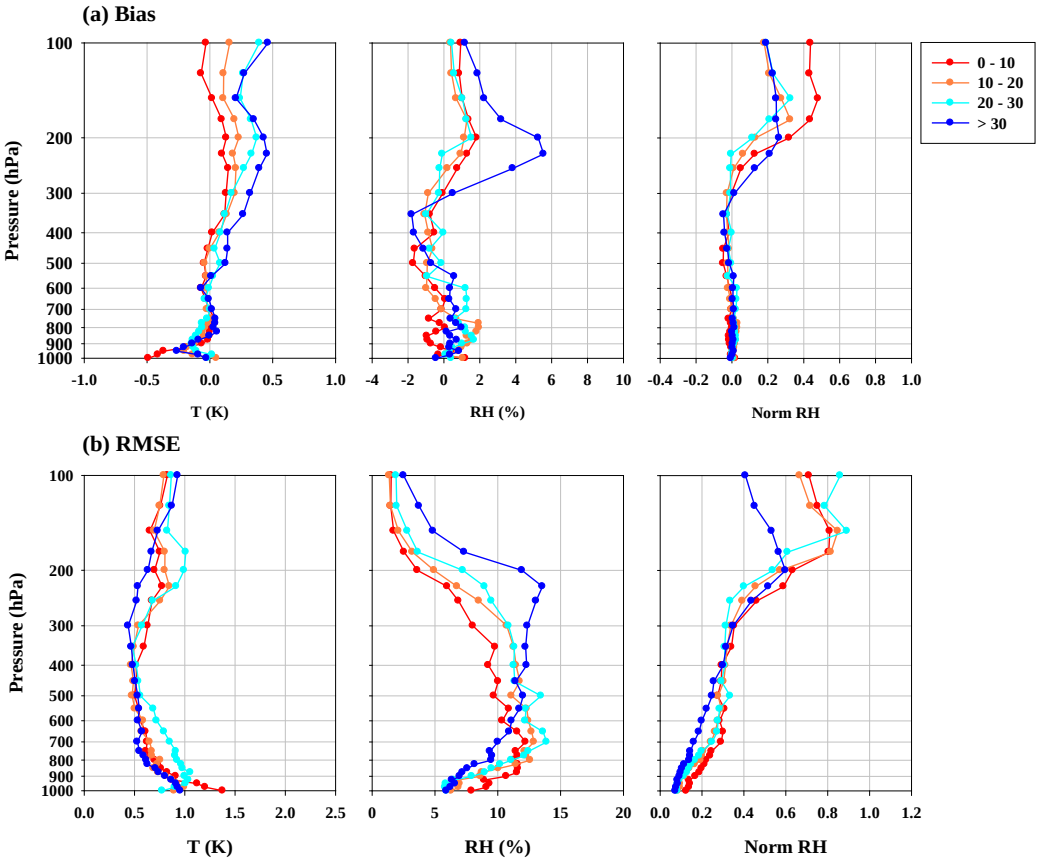

Figure 8. Error statistics of the United Kingdom Met Office model (UM-UKMO) temperature and relative humidity profiles classified according to the total precipitable water between the surface and 100 hPa (TPW, in kg m^{-2}).

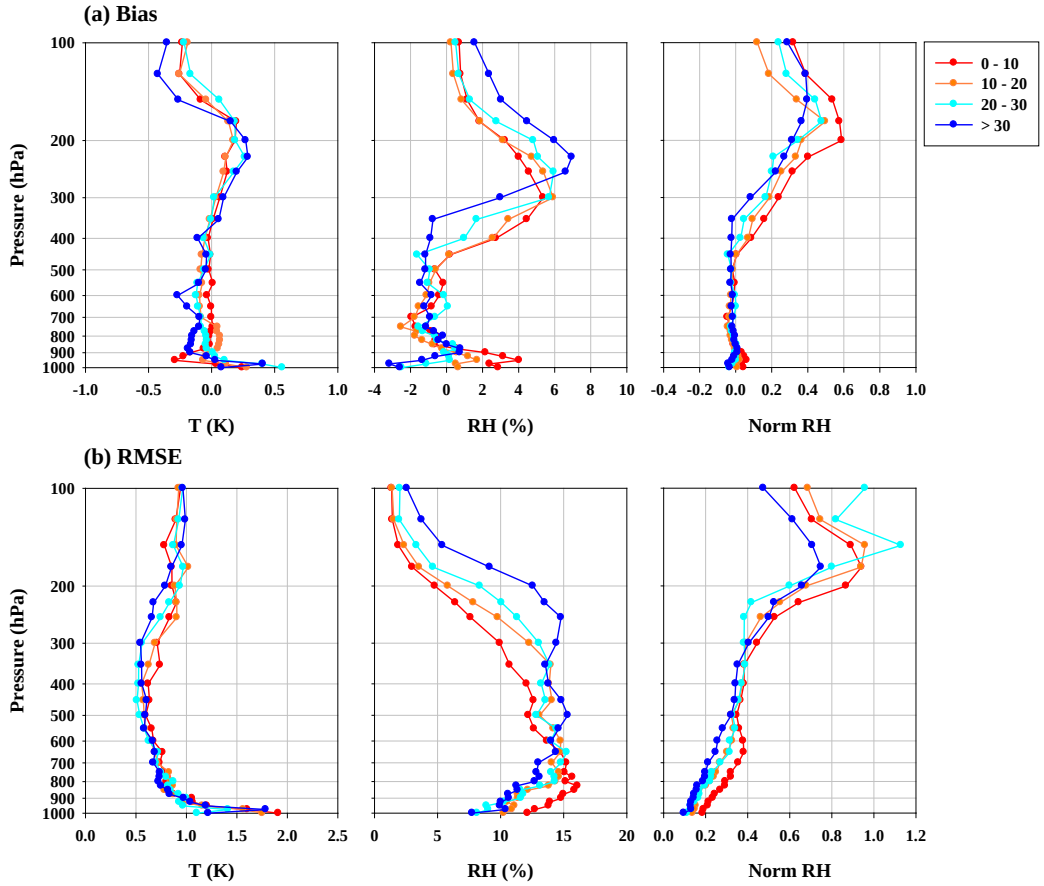

Figure 9. Error statistics of the European Centre for Medium-Range Weather Forecasts Interim reanalysis (ERA-Interim) temperature and relative humidity profiles classified according to the total precipitable water between the surface and 100 hPa (TPW, in kg m^{-2}).

4. IASI channel selection

Here we provide a methodology of how the IASI channel selection is made in this study. Because the information of key atmospheric variables provided by channels can be changed under certain atmospheric condition (e.g., overcast condition), the selected channels may be also changed depending on imposed specific conditions (Migliorini, 2015). In this study, we attempt to select IASI channels under the clear-sky oceanic condition.

To update the IASI channels used in the UM data assimilation process, the channel selection is performed, which is based on the statistical approach using 1D-Var analysis. The channels are selected from the preselected 314 channels (i.e. 300 channels suggested by Collard 2007 and 14 channels for monitoring), because these channel data are officially distribute by GTS system and now available in the UM assimilation system.

4.1. Basic framework for IASI channel selection

4.1.1. One-dimensional variational analysis

The IASI channel selection was made after quantifying the degree of improvement of atmospheric variables retrieved from 1D-Var. In this procedure, 1D-Var minimizes the cost function $J(\mathbf{x})$ given as follows:

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}_{\mathrm{o}}\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}_{\mathrm{o}}\right) + \frac{1}{2}\left[\mathbf{y}_{\mathrm{o}} - \mathrm{H}\left(\mathbf{x}\right)\right]^{\mathrm{T}} \mathbf{R}^{-1}\left[\mathbf{y}_{\mathrm{o}} - \mathrm{H}\left(\mathbf{x}\right)\right] \quad (1)$$

where $\mathbf{x}$ is the control variable consisting of atmospheric variables, $\mathbf{x}_{\mathrm{b}}$ is the background variable, and $\mathbf{y}_{\mathrm{o}}$ is the observed radiance at the given sensor channels. $\mathbf{B}$ is the background-error covariance matrix associated with background variable $\mathbf{x}_{\mathrm{b}}$. $\mathbf{R}$ is the observation-error covariance matrix that is the sum of the forward model error and instrumental error. The superscript 'T' indicates the matrix transpose and H($\mathbf{x}$) is the forward-modeled radiance associated with the control variable $\mathbf{x}$ (Rodgers, 2000; Deblonde et al., 2007; Pavelin et al., 2008).

The optimal solution is found by minimizing Eq. (1) or by solving it gradient equation. The gradient equation with respect to x is derived as

$$\nabla_{x} J = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}_{\mathrm{o}}\right) - \mathbf{H}^{\mathrm{T}}\left(\mathbf{x}\right)\mathbf{R}^{-1}\left[\mathbf{y}_{\mathrm{o}} - \mathrm{H}\left(\mathbf{x}\right)\right] \quad (2)$$

where $\mathbf{H}^{\mathrm{T}}$ is the adjoint operator of the Jacobian matrix, which contains the derivatives of each measurement with respect to each element of the state vector. $\mathbf{H}$ is the tangent linear function of H at point x (where H here is the potentially nonlinear observation operator). In general, there is no general solution to this

equation. In order to solve this equation, we only consider the linear case, i.e. when

$$\mathbf{H}(\mathbf{x}) = \mathbf{H}(\mathbf{x}_o) = \mathbf{H} = \text{ constant, } \mathrm{H}(\mathbf{x}) = \mathrm{H}(\mathbf{x}_o) + \mathbf{H}(\mathbf{x} - \mathbf{x}_o) \qquad (3)$$

The optimal solution equation is obtained by setting the gradient of the cost function zero and using Eq. (3) is expressed as follows:

$$\mathbf{x} = \mathbf{x}_o + \left[\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right]^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\left[\mathbf{y}_o - \mathrm{H}(\mathbf{x}_o)\right] \qquad (4)$$

Matrix manipulation makes an equivalent equation which is often computationally more efficient:

$$\mathbf{x} = \mathbf{x}_o + \mathbf{B}\mathbf{H}^{\mathrm{T}}\left[\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right]^{-1}\left[\mathbf{y}_o - \mathrm{H}(\mathbf{x}_o)\right] \qquad (5)$$

The aim of the 1D-Var scheme is to minimize the cost function by solving its gradient equation. Thus, the minimum cost function is found by applying Newtonian iteration method. An optimal solution to Eq. (4) is expressed as follow:

$$\mathbf{x}_{n+1} = \mathbf{x}_o + \mathbf{K}_1\left[\mathbf{y}_o - \mathrm{H}(\mathbf{x}_n) - \mathbf{H}(\mathbf{x}_n)(\mathbf{x}_o - \mathbf{x}_n)\right]$$
$$\mathbf{K}_1 = \left[\mathbf{H}^{\mathrm{T}}(\mathrm{x}_n)\mathbf{R}^{-1}\mathbf{H}(\mathrm{x}_n) + \mathbf{B}^{-1}\right]^{-1}\mathbf{H}^{\mathrm{T}}(\mathrm{x}_n)\mathbf{R}^{-1} \quad (6)$$

if the number of measurements is more than that of unknowns. Alternative optimal solution to Eq. (5) is expressed as follow:

$$\mathbf{x}_{n+1} = \mathbf{x}_o + \mathbf{K}_2\left[\mathbf{y}_o - \mathrm{H}(\mathbf{x}_n) - \mathbf{H}(\mathbf{x}_n)(\mathbf{x}_o - \mathbf{x}_n)\right]$$
$$\mathbf{K}_2 = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{x}_n)\left[\mathbf{H}(\mathbf{x}_n)\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{x}_n) + \mathbf{R}\right]^{-1} \quad (7)$$

if the number of unknowns is more than that of measurements.

4.1.2. Simulation of IASI brightness temperature and background

In order to perform 1D-Var retrievals, observed radiances for the selected IASI channels and background atmospheric variables are required as input data. Because this study quantitatively assesses the impact of inclusion of each of the pre-selected 314 IASI channels in terms of the degree of improvement of control variables obtained from 1D-Var, it is necessary to validate the 1D-Var outputs of temperature and humidity profiles against collocated in-situ measurements, such as radiosonde data. However the assessment of the improved performance of the IASI channels used in the 1D-Var retrievals assumes that the true atmospheric

conditions are known. The observed satellite radiances and background data are then theoretically generated with the given true atmospheric conditions (shown in Fig. 10).

For reference, we use a set of 600 profiles, which were randomly selected under clear-sky oceanic conditions from the European Centre for Medium-range Weather Forecasts (ECMWF) Re-Analysis Interim (ERA-Interim) data in year 2012. The reference data comprise temperature and moisture profiles at 37 standard pressure levels, surface pressure, surface humidity, and skin temperature. From these data, the IASI infrared brightness temperatures and background profiles are obtained and used as input data for the 1D-Var analysis in the NWP SAF (NWP satellite application facility) 1D-Var package version 3.4 (Weston et al., 2013).

For the IASI brightness temperature simulation, we use the Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV) model, version 10, a fast radiative transfer model for data assimilation and retrieval systems (Matricardi et al., 2004). After interpolating temperature and water vapor profiles of the reference data to the required 43 standard pressure levels for RTTOV, IASI brightness temperatures at the preselected 314 channels (Collard's 300 plus 14 monitoring channels) are simulated with clear-sky oceanic condition. 43 levels are chosen because the UM 1D-Var uses the same 43 levels. In this simulation, random

measurement errors are added to the simulated brightness temperatures, which are derived from the IASI observation-error covariance matrix used for the operational data assimilation system of the UM. Thus, observation errors in this simulation are thought to satisfy an unbiased Gaussian distribution and a diagonal observation-error covariance matrix is employed, implying that observation errors between any two IASI channels are assumed to be uncorrelated. However, the use of the diagonal observation-error covariance matrix may be problematic because of correlations existing between different channels (Bormann and Bauer, 2010; Weston et al., 2014). These correlations may arise from a variety of sources such as instrument noise errors, forward model errors, representativeness errors, and systematic errors related to the pre-processing. Biggest contributions are from representativeness errors and systematic errors. Nevertheless, considering that theoretically-generated observations and background data in this study are independent of the scale mismatch between the observation and model, the representativeness errors caused by the mismatch can be ignored. The systematic errors can be ignored as well, because we use simulated radiances not requiring the quality-control step. On the other hand, observation-error correlations between neighboring channels caused by apodization were already taken into account because Collard's 300 channels were selected by removing nearest neighboring channels. However, two channels (689.50 cm^{-1} and 705.25 cm^{-1}) among 14 CNES

monitoring channels are next to Collard channels (689.75 cm^{-1} and 705.50 cm^{-1}), respectively. Nevertheless, the error correlation between CNES channels and neighboring Collard channels can be ignored, because random observation errors derived from the diagonal observation-error covariance matrix were added to the simulated IASI brightness temperatures. Therefore, it is reasonable to employ the diagonal observation-error covariance matrix with instrument noise errors so that IASI simulated radiances are suitably weighted against the background for 1D-Var.

Background estimates of the atmospheric state are also required for the 1D-Var simulation. In the operational assimilation system of the UM, 6-hour forecasts from the previous data assimilation cycle are used as a background. However, in our work, the background is generated by adding random forecast errors to the reference data. The random forecast errors are calculated from the IASI background-error covariance used by the operational UM system. The background-error covariance matrix consists of the error covariance between 73 atmospheric and surface parameters (i.e., temperature at 43 pressure levels from 1013.25–0.1 hPa, water vapor at 26 pressure levels from 1013.25–122 hPa, surface temperature, surface humidity, skin temperature, and surface pressure). The forecast errors are then expressed by:

$$\mathbf{x}_{\mathrm{o}} = \mathbf{x}_{\mathrm{t}} + \sum_{i} \varepsilon_i \lambda_i^{1/2} \mathbf{E}_i \tag{8}$$

where $\mathbf{x}_0$ is the perturbed background profile, $\mathbf{x}_t$ is the reference profile, $\mathbf{E}_i$ and λ_i are the eigenvector and eigenvalue of the background-error covariance matrix, respectively, and ε_i is a random number with a zero mean and a unit standard deviation (Pavelin et al., 2008). For this work, we adopted the operationally used background-error covariance matrix, which is derived from the so-called National Meteorological Center (NMC) method (Parrish and Derber, 1992; English, 1999).

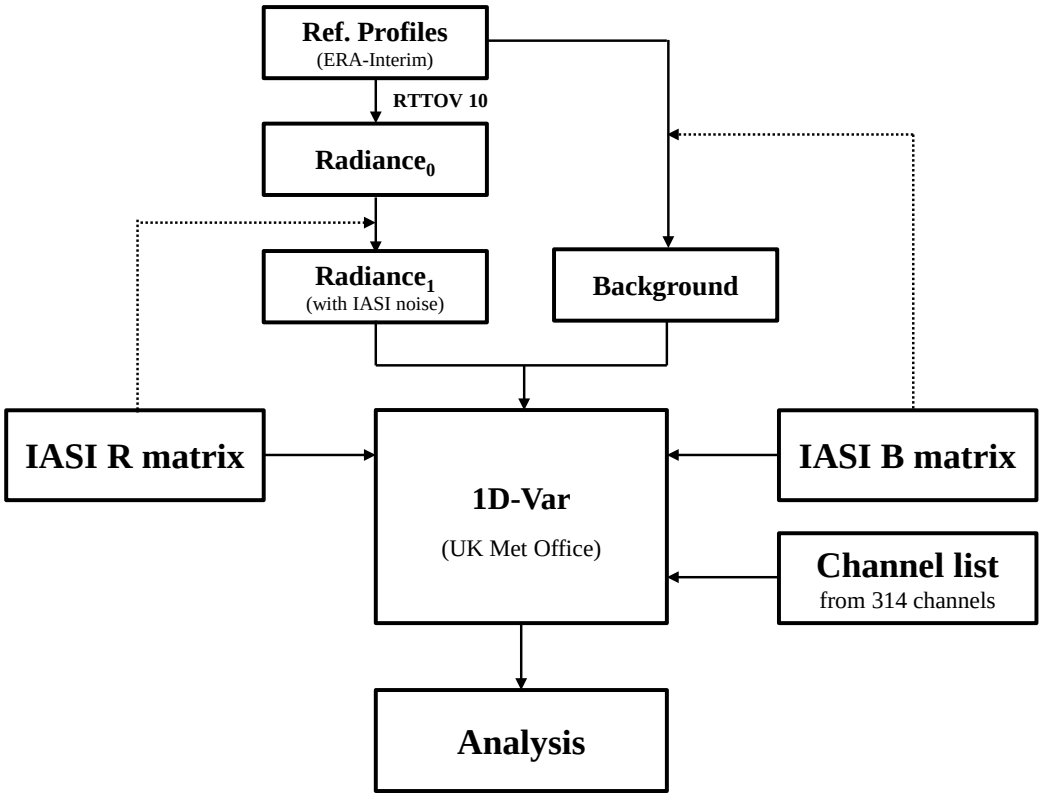

Figure 10. Schematic diagram showing the basic framework of IASI channel selection using the 1D-Var scheme. B and R are background-error covariance and observation-error covariance, respectively.

4.1.3. Method for evaluating the impact of IASI channels

The pre-defined information contents such as the degree of freedom for signal (DFS) and the entropy reduction (ER), which were employed in Collard (2007), are calculated using the full error-covariance matrices of both background and analysis, and they represent the degree of improvement of the analysis error covariance against the background error covariance. However, because full error covariance is required for estimating such information contents, it is difficult to evaluate the direct impact of channels on the retrieved variables when 1D-Var is employed. Here, since we aim to select channels based upon the direct contribution to the accuracy of retrieved atmospheric variables using 1D-Var, the channel selection in this work is based on a statistical method involving the comparison of 1D-Var results with background data.

Since in this case the true values of the atmospheric parameter xt are known, the root-mean-square errors (RMSEs) of the analysis and background profile can be defined as follows:

$$\mathrm{RMSE}(\mathrm{A})_{i,j} = \sqrt{\frac{\sum_{k=1}^{M}\left(x_{i,j,k}^{a} - x_{j,k}^{t}\right)^{2}}{M}}, \quad \mathrm{RMSE}(\mathrm{B})_{j} = \sqrt{\frac{\sum_{k=1}^{M}\left(x_{j,k}^{b} - x_{j,k}^{t}\right)^{2}}{M}} \qquad (9)$$

In Eq. (9), subscript *i* denotes the IASI channel number (i.e., *i* = 1–314) used, subscript *j* denotes the atmospheric and surface parameter number (i.e., *j* = 1–43 for temperature at 43 pressure levels from 1013.25–0.1 hPa, *j* = 44–69 for water vapor at 26 pressure levels from 1013.25–122 hPa, and *j* = 70–73 for surface 2m temperature, surface 2m humidity, skin temperature, and surface pressure); subscript *k* denotes an index for the 600 reference profiles (i.e., *M* = 600, *k* = 1–*M*). The superscripts a, b, and t represent the analysis, background, and true reference profile, respectively. Thus, $\mathrm{RMSE(A)}_{i,j}$ is the RMSE between the analysis and reference for the j^{th} analysis parameter, for an *i* number of IASI channels selected for 1D-Var. $\mathrm{RMSE(B)}_j$ is the RMSE between the background and reference for the j^{th} parameter. To select the total number of *i* channels out of the 314 provisional channels, instead of random selection, we devised an efficient method described in Section 4.1.4.

In order to assess the overall impact of a particular set of selected IASI channels on the improvement of the analysis parameters for 1D-Var, the following channel score index (CSI) is defined as a figure of merit:

$$\mathrm{CSI} = \sum_{j=1}^{N} \left[1 - \frac{\mathrm{RMSE}\left(\mathrm{A}\right)_{i,j}}{\mathrm{RMSE}\left(\mathrm{B}\right)_{j}} \right] W_j \qquad (10)$$

where w_j is the weight associated with the j^{th} atmospheric parameter and N is the total number of atmospheric parameters used for calculating the CSI (i.e., $N = 73$). The CSI value indicates the degree of improvement in the analysis profile over the background profile. A positive CSI means that the analysis error is smaller than that of the background, implying that the selected IASI channels for the 1D-Var analysis are contributing positively.

Concerning w_j for water vapor in Eq. (10), it is noted that the contribution of upper-tropospheric water vapor to the CSI tends to be comparatively large and highly variable despite small amounts of water vapor (not shown). In order to show the water vapor background error varying with height, the background error normalized by level's mean value is plotted in Figure 11a. The normalized error becomes close to the level's value itself above the 400 hPa level. Because of the larger error percentage in the upper layers, any small RMSE improvement in the upper layer contributes to the CSI with a much larger ratio, compared with the same RMSE improvement in the lower layers. Therefore, it is necessary to apply weights to the water vapor contribution to the CSI, in order to alleviate the comparatively large influence from the upper-tropospheric water vapor. The weight for the water vapor at each level is obtained by using the mean profile obtained from 600 samples profiles; the weight at a particular level is estimated by normalizing the mean water vapor mixing ratio at that level with the mean water

vapor at the lowest level (Fig. 11b). The weights (for the 44th to 69th water vapor parameters) at 26 pressure levels decrease gradually from 1 at the surface to a much smaller value at the top 122 hPa level. Weights for other atmospheric parameters are set to 1.

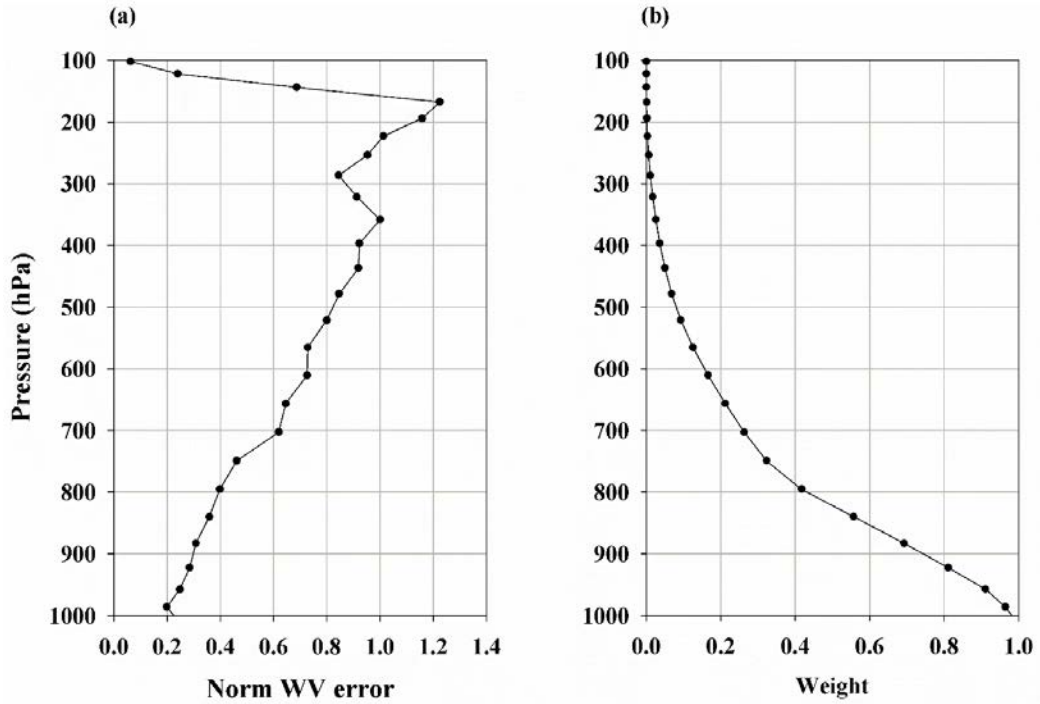

Figure 11. (a) The normalized background water vapor error against the mean water vapor mixing ratio (kg/kg) and (b) vertical structure of the weight applied to the water vapor contribution to the channel score index (CSI).

4.1.4. Channel selection

First, the CSI is calculated for each of the provisional 314 IASI channels, in order to find the channel that has the largest contribution to the CSI. Then, that channel is considered to be the first selected channel. Second, each of the remaining 313 channels is added to the first selected channel (forming two channels) so that the impact of the two selected channels on the CSI can be estimated. The channel that gives the largest contribution to the CSI after it is combined with the first selected channel is considered to be the second selected channel. The third channel is selected by determining which channel of the remaining 312 channels gives the largest contribution to the CSI when added to the first two selected channels. This procedure is repeated until the last channel is left.

4.2. Characteristics of new IASI channels

Figure 12 shows the CSI calculated from the 1D-Var simulation with the selected channels by applying the procedures described in Section 4.1. It is shown that the CSI increases rapidly with an increasing number of selected channels, reaching up to 85% of the total CSI at 60 selected channels and 93% at 100. The slope tends to decrease slowly after 100 selected channels, and the CSI approaches

an asymptotic value, especially when the number of selected channels becomes close to 200. After 200 selected channels, additional channels have little impact on the CSI, suggesting that the first 200 channels contain most of the information required for improving the background atmospheric state. Furthermore, there is a tendency for less convergence in 1D-Var when more than 200 channels are used. Considering that IASI observations causing convergence failure in 1D-Var should be discarded before further use for the data assimilation, the first 200 channels yielding most of the CSI could be considered optimal for the IASI data assimilation for atmospheric temperature and water vapor retrievals.

It can also be seen that the selected 200 channels are distributed largely over CO_2 (650–750 cm^{-1}), H_2O (1350–2000 cm^{-1}), and O_3 (1030 cm^{-1}) absorption bands in the IASI spectrum from 645–2760 cm^{-1} (Fig. 13). Most of the channels over the CO_2 absorption band recommended by Collard (2007) were selected in this new selection (Fig. 13b). Selected H_2O channels seem to be divided into two groups, one near 2000 cm^{-1} and the other over a 1300–1600 cm^{-1} band. Since the former carries the water vapor information over the lower troposphere and the latter over the mid to upper troposphere, H_2O channels near 1600 cm^{-1} sensitive to the upper-tropospheric water vapor are very rare. The fewer selected H_2O channels for the upper-troposphere, in comparison to the near full selection of available CO_2 channels, are likely due to the larger instrumental noises of H_2O channels, which

can cause a larger observation-error covariance. Instrumental noise errors for H_2O channels appear to be larger and more variable than those for CO_2 channels – see Figure 2 of Collard (2007). Given the fact that the error covariance is closely related to the instrumental noise, H_2O channels are less capable of increasing the CSI and thus are selected less in comparison to the CO_2 channels. Beside the major selection of channels over CO_2 and H_2O absorption bands, selected channels are found over the 1030 cm^{-1} O_3 absorption band and in the split window channels over 900 cm^{-1} and 1100 cm^{-1} spectral bands.

It is of interest to examine how the 200 newly selected channels are different from the 183 IASI channels, which have been used operationally in the UM global forecast system at the UKMO since year 2007 (Weston, 2011). This is of interest because the UM model has been improved since IASI data was first used.

It is found that 149 out of the operationally used 183 channels are also shown in the new 200 channels (Fig. 14a) and thus 34 channels (Fig. 14b) are not included in the 200 channels. Fifty-one channels selected only from the 200 new channels (Fig. 14c) are located mainly in the O_3 band over 1030 cm^{-1}, the H_2O bands around 1350 cm^{-1} and 2000 cm^{-1}, and in band 3 (2000–2760 cm^{-1}). From the comparison of the weighting functions between the newly selected 200 channels and the operational 183 channels, the vertical distributions of weighting functions for new channels are found to be more evenly distributed (not shown). It seems due to the

addition of more middle- and upper-tropospheric channels while removing surface-sensitive channels.

It is not surprising to find channels over the O_3 band because the UM OPS noted that the use of O_3 channels in the IASI data assimilation in 2007 could cause a divergence during the 1D-Var minimization process. High-peaking H_2O channels (1300–1600 cm^{-1}) whose temperature Jacobian peaks are located above 520 hPa were also removed because of the same minimization problem (Hilton et al., 2009). However, six high-peaking H_2O channels (corresponding to channel wave numbers 1436.75, 1456.75, 1521.25, 1539.0, 1540.25, and 1542.0 cm^{-1}) survived for the assimilation. These channels have Jacobians without well-defined peaks and with long tails into the stratosphere, but have a slightly positive influence on the forecasts (Smith, 2014). In our new H_2O channel selection, the channels that survived were replaced by other H_2O channels with high peaks above 520 hPa.

In addition, the IASI band 3 channels have not been utilized because these channels also have higher instrument noise as well as water vapor channels, but are mainly at risk for being contaminated by solar radiation reflected from surface during the day (Hilton at al., 2009). However, the IASI band 3 channels show the significant impact on the accuracy of analysis from 1D-Var, meaning that most of these channels are selected below one hundred of accumulative selected channels.

Thus, even though it is still difficult for the IASI band 3 channels to be used in the daytime because of sun-glint, the use of these channels in the night time can contribute to the performance of data assimilation.

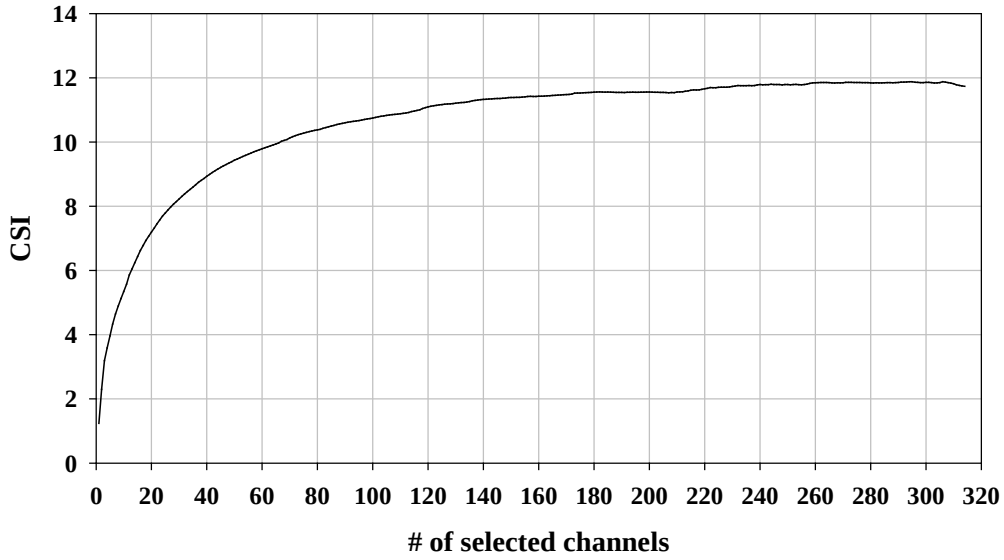

Figure 12. Channel score index (CSI) depending on the number of selected IASI channels.

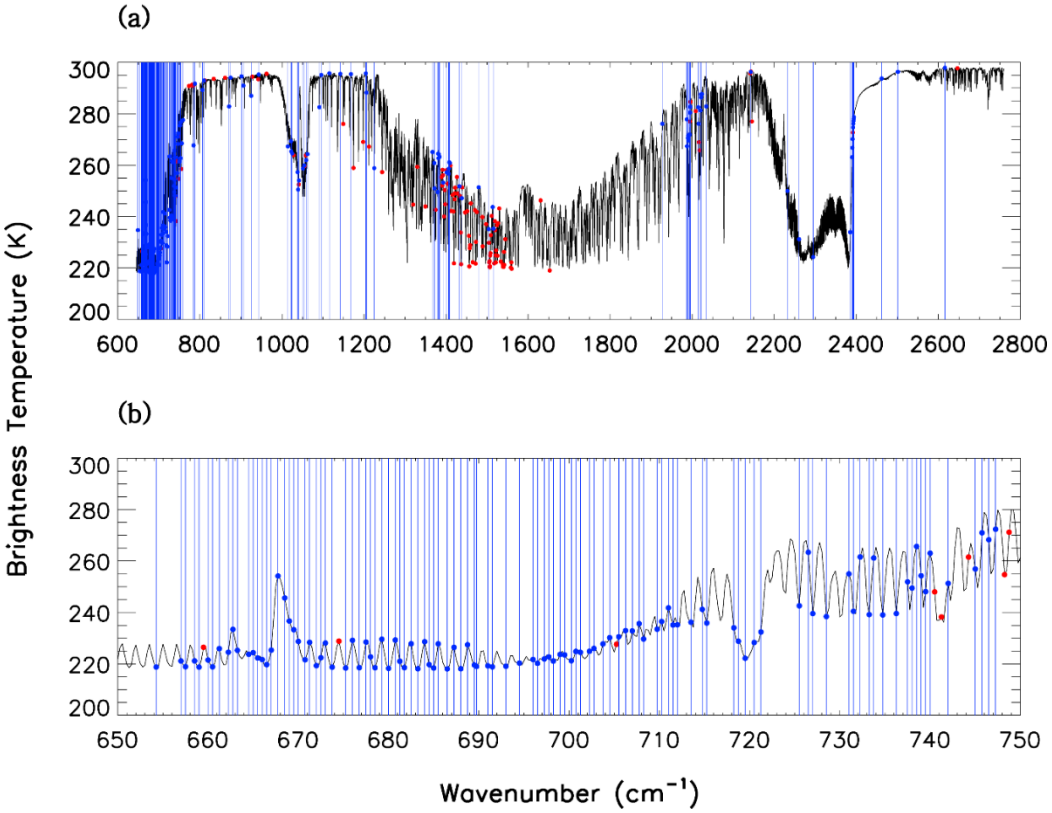

Figure 13. Spectral distributions of the 200 new selected IASI channels over (a) infrared band (600–2800 cm^{-1}) and (b) CO_2 absorption band (650–750 cm^{-1}). Blue bars and red dots indicate the 200 selected IASI channels and remaining 114 channels among the provisional 314 IASI channels, respectively.

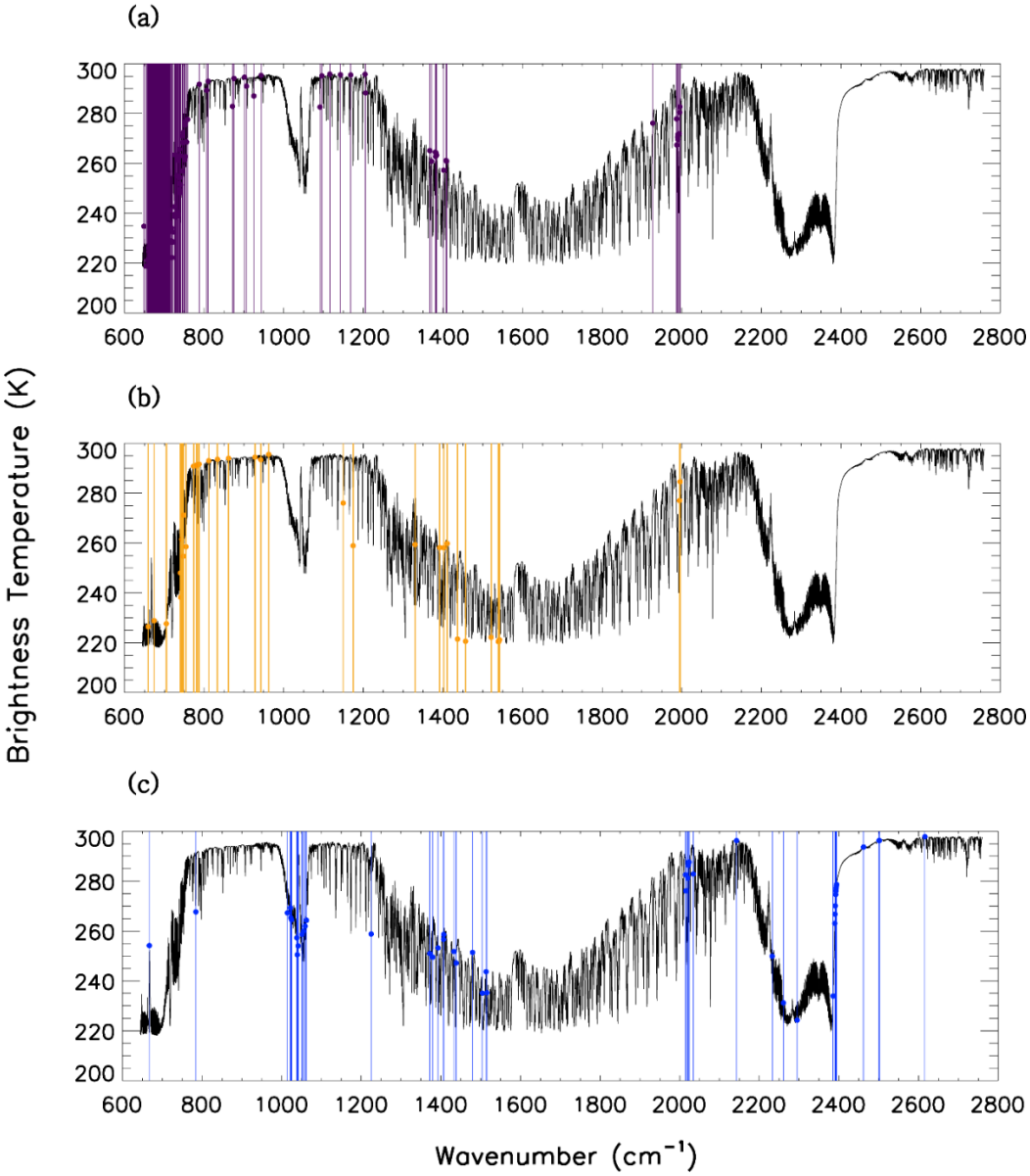

Figure 14. Spectral distributions of (a) 149 channels commonly shown in both operational 183 channels and new 200 channels, (b) 34 operationally used channels excluded from the new 200 channels, and (c) 51 new channels excluded from the operational 183 channels.

4.3. Comparison with the predefined selection method

Since the CSI is estimated from errors of 1D-Var retrieved and background atmospheric variables, it is interesting to examine how the use of full error covariance might influence on the channel selection. In this, the ER method is employed, which can be estimated from the error-covariance matrix reflecting error correlations between atmospheric parameters at different levels (Rodgers, 2000), i.e.:

$$\mathrm{ER} = -\frac{1}{2} ln \det\left(\mathbf{A}\mathbf{B}^{-1}\right) \tag{11}$$

where **A** and **B** indicate the analysis and background error covariance matrices. In this ER method, selection procedures are same as in the CSI approach except for the use of ER. In order to select the optimal channels for a diverse range of atmospheric profiles, an ER mean is calculated from same 600 atmospheric profiles used for the CSI approach.

$$\overline{\mathrm{ER}_i} = \frac{\sum_{j}^{M}\left[-\frac{1}{2} ln \det\left(\mathbf{A}_{i,j}\mathbf{B}^{-1}\right)\right]}{M} \tag{12}$$

where subscribe i denotes the channel number (i.e., i = 1–314) used for the 1D-Var and subscribe j the index for 600 atmospheric profiles (i.e., M = 600, j = 1–M). Thus, $\overline{\mathrm{ER}_i}$ indicates an ER mean from i number of IASI channels used for 1D-Var.

Figure 15 shows the ER mean at given selected IASI channels. Compared to CSI results (Fig. 12), the ER increases rather slowly, reaching to 82% at 100 and 90% at 170 selected channels – compare with CSI results of 93% at 100 and 96% at 170. Furthermore, the ER steadily increases even after 200, where the ER reaches about 94% of the total. These results suggest that the selected channels may be close to 314 when the ER method is used, much more channels than suggested by the CSI method.

The spectral distribution of the first 200 channels selected with the ER is given for the comparison with CSI results – Figure 16. Compared with CSI results (Fig. 13a), the ER method selected less channels over CO_2 absorption band but more H_2O channels over 1350–2000 cm^{-1}. In particular, much more upper-tropospheric H_2O channels over 1350–1600 cm^{-1} were selected. Substantial difference in the two sets of selected channels suggests that each channel contributes differently to two indices.

In order to examine how each channel contributes to the ER and the CSI, we estimate the ER and the CSI at each channel of 314 provisional channels – results

are given in Figure 17. As expected, channels over CO_2 (650–750 cm^{-1}), H_2O (1350–2000 cm^{-1}) absorption bands, and window regions (900 cm^{-1} and 1100 cm^{-1} spectral bands) contribute to the ER through the improvement of temperature, water vapor, and surface variables, respectively (Fig. 17a). It is noted that contributions of water vapor to the ER for H_2O channels, particularly associated with upper tropospheric humidity, are much stronger than for channels at any other spectral bands. This suggests that upper-tropospheric H_2O channels are selected first, prior to other channels, as shown in Figure 16. Heavy contributions by water vapor in the ER approach may be alleviated if weights are given to water vapor profiles. However, it is difficult to give weights while counting all the correlations between levels. Therefore, even if the 1D-Var approach is adopted for the ER method, the procedures should be taken individually over separate absorption bands.

The CSI results given in Figure 17b show similar features found from the ER (Fig. 17a). But, the contribution by water vapor to the CSI over the H_2O absorption band appears small due to the applied weights.

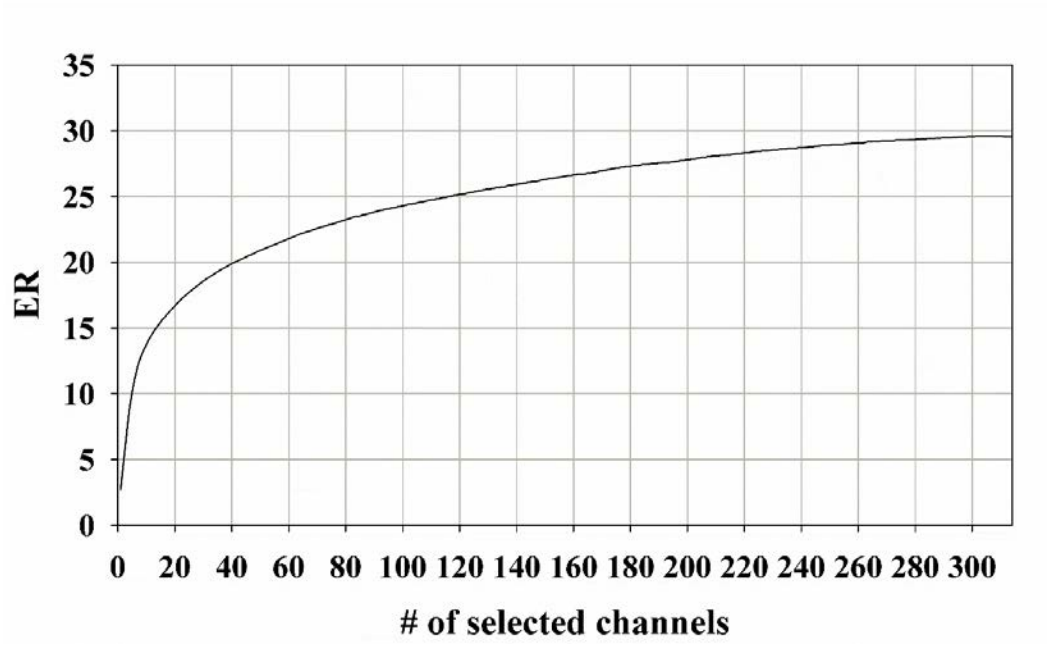

Figure 15. Entropy reduction (ER) with the number of selected IASI channels.

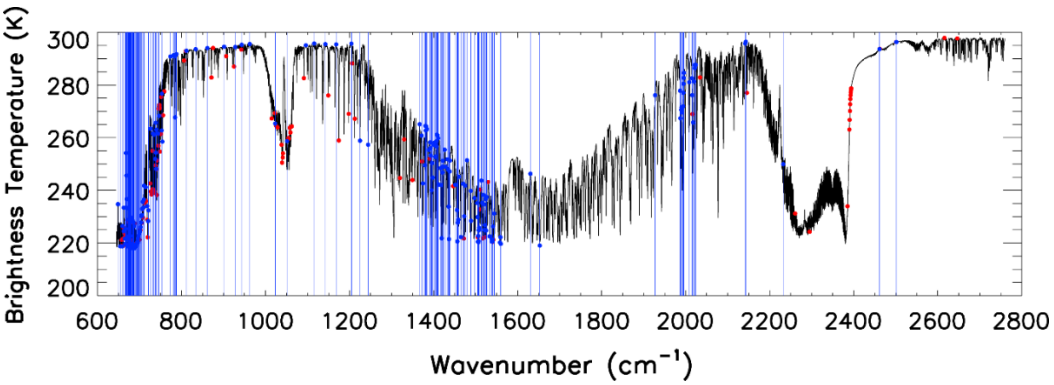

Figure 16. Spectral distribution of the 200 IASI channels with entropy reduction (ER) over infrared band (600–2800 cm^{-1}). Blue bars and red dots indicate the 200 selected IASI channels and remaining 114 channels among the provisional 314 IASI channels, respectively.

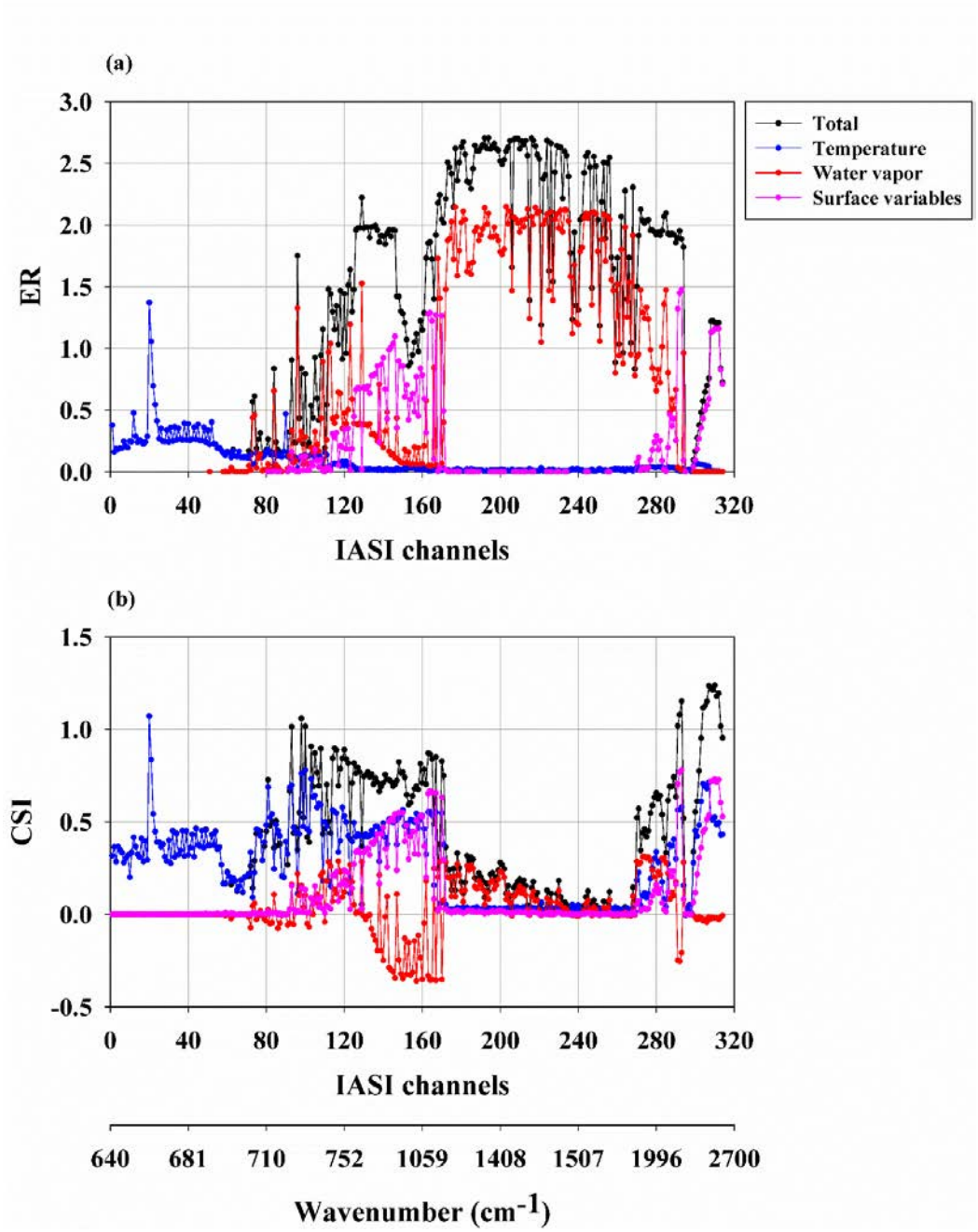

Figure 17. Contribution of temperature (blue), water vapor (red), and surface variables (pink) to (a) the total entropy reduction (ER) and (b) the channel score index (CSI) for each of IASI 314 channels.

4.4. IASI channel selection from EUMETSAT 500 channels

In this study, we tried to select optimal channels from 314 provisional IASI channels that have been operationally received via EUMETCAST in the UM assimilation system. It would have been much desirable if the selection was made from the total 8461 IASI channels using the proposed CSI method. However, the computational burden caused by1D-Var runs appears formidable; even in the case of selecting 200 channels from 314 pre-selected channels, the total number of 1D-Var runs required is 25,740,000 (= (314+313+312+ ··· +115) × 600 profiles). Being aware of the limitation brought by starting with a small number of pre-selected channels, it is of interest to examine how the channel selection by the CSI method is influenced by a different set of pre-selected channels. It is of further interest to examine how the new set of channels, if there is any, is comparatively informative against the old set (here, for example, EUMETSAT 314 channels vs. 500 channels). As a case of point, we give an effort of selecting new channels from pre-selected EUMETSAT 500 channels (distributed from 15 October 2014), by using the same procedures provided in this paper.

EUMETSAT 500 channels comprise 300 channels suggested by Collard (2007), 43 channels chosen by Collard and McNally (2009), 134 channels from Martinet et al. (2014) work, and 23 monitoring channels. Using the same CSI selection procedure using 1D-Var taken in this study, first 220 channels were

selected from EUMETSAT 500 channels because the cumulative CSI tends to level off after 180 (Fig. 18). In other words, additional channels after first 180 channels have little impact on the cumulative CSI. Therefore, we determined the first 200 channels to be optimal and these are referred to as EUMETSAT 200.

Compared with 200 channels selected from 314 channels (referred to as Collard 200), EUMETSAT 200 channels include more split window channels and low-tropospheric water vapor channels from Martinet's 134 channels and some additional channels from 43 CO_2 channels of Collard and McNally (2009) – Figure 19. In spite of those changes in spectral locations, however, numbers of selected channels at each spectral band appear to be nearly same (Table 3). It means that some channels of Collard 200 channels at each spectral band are replaced by channels in Martinet's 134 channels and Collard's 43 channels, but adding/removing little in the given spectral band. Nonetheless, from the perspective of total cumulative CSI, EUMETSAT 200 channels yield the CSI 6.7% larger than from Collard 200 channels, suggesting that EUMETSAT 200 channels are more informative than Collard 200 channels.

This results suggest that it may be necessary to select a new subset of channels if the environment is changed, in association with NWP model or supply of satellite data, in order to perform more efficient data assimilation. Considering that all channels across the absorption bands should not be equal, determining

most important channels in sequence must be a valuable effort.

Table 3. Number of selected channels depending on spectral ranges.

Band	**wavenumber (cm^{-1})**	**200 channels from 314 channels**	**200 channels from 500 channels**
A	600–750	110	112
B	750–1200	32	31
C	1200–2000	32	35
D	2000–2700	26	22
Total		200	200

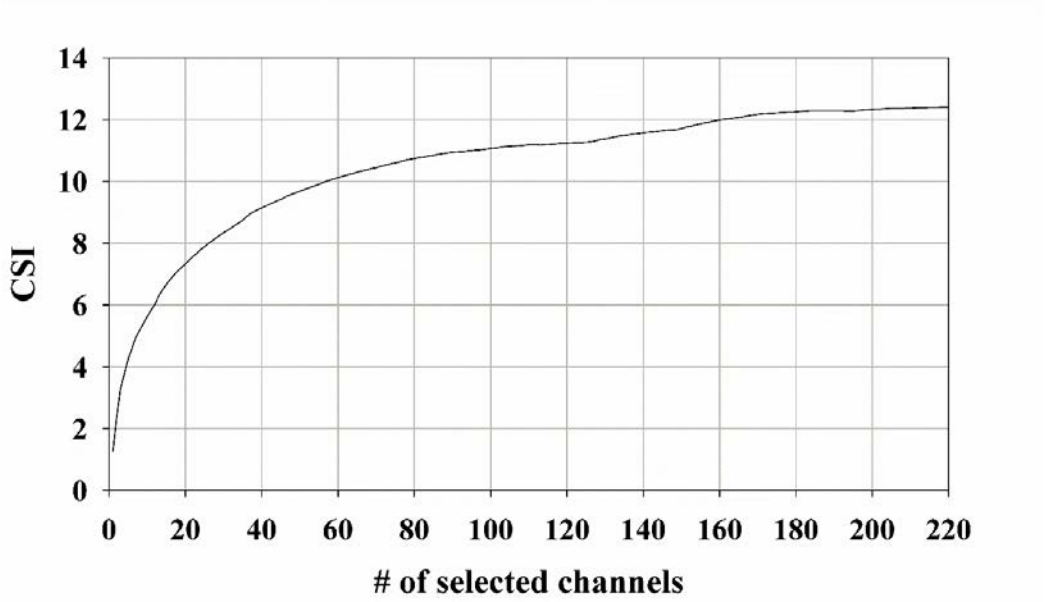

Figure 18. Channel score index (CSI) depending on the number of selected IASI channels from EUMETSAT 500 channels.

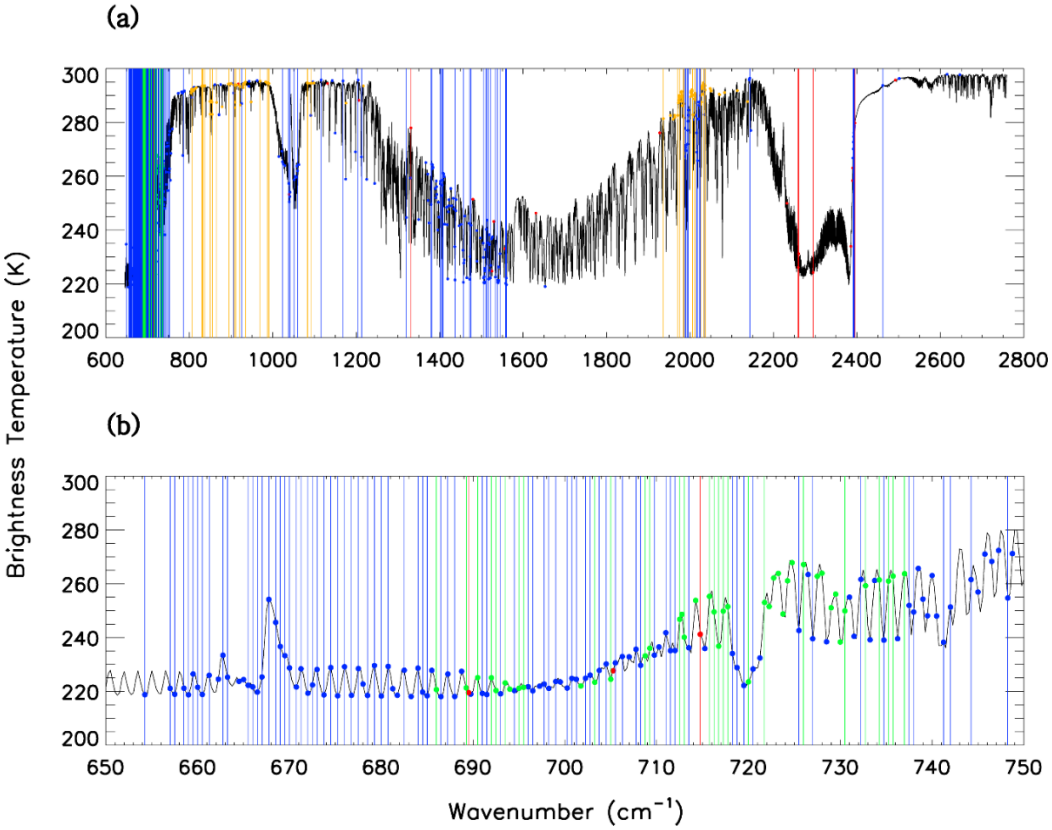

Figure 19. Spectral distribution of 200 IASI channels selected from EUMETSAT 500 channels over (a) infrared band (600–2800 cm^{-1}) and (b) CO_2 absorption band (650–750 cm^{-1}). Vertical lines in color indicate the selected channels while dots represent channels not in 200 channels. EUMETSAT 500 channels consist of Collard's 300 channels (blue), Collard and McNally's 43 channels (green), Martinet's 134 channels (yellow), and 23 monitoring channels (red).

4.5. Impact of new IASI channels on 1D-Var analysis

In order to validate the impact of new selected IASI 200 channels on atmospheric parameter retrieved from 1D-Var simulation, the error analysis (i.e. RMSE) is performed using the analysis data resulting from 1D-Var simulation in Figure 20. Since this study is a kind of simulation study meaning that the radiances of IASI channels and background data, as input data for 1D-Var simulation, are artificially generated from reference profiles, the error of analysis and background could be simply calculated with about six thousands profiles that are randomly extracted from global ERAI reanalysis data.

As shown in Figure 20a, the RMSEs of temperature profiles of both the operational IASI channels and new IASI channels are smaller than that of background. However, the new IASI channels show the lower RMSEs in the lower troposphere than that of operational IASI channels, and the improvement of temperature in the upper troposphere is also found though these values are slightly small. The enhancement of temperature profile is likely to be induced by the added channels of IASI band 3 in Figure 14c, because most of CO_2 channels in IASI 314 candidate channels are included in the operational and new IASI channels. In particular, the spectral band near a wavenumber of 2300 cm^{-1} is well known as CO_2 absorption band at a wavelength of 4.3 μm (Susskind et al., 2011).

The RMSEs of water vapor profiles show the similar pattern with the

validation results of temperature profile. The RMSEs of the operational and new IASI channels represent the better performance in compared with that of background, but the errors of water vapor from the new IASI channels is smaller than that of operational IASI channels over the whole troposphere. As shown in Figure 20b, although the degree of improvement of water vapor in the upper troposphere seems to be shown insignificantly, the slight improvement of water vapor could play an important role in understanding the upper atmospheric state, as a small amount of water vapor in the upper troposphere can influence on the energy balance and climate system in the Earth system (Dessler and Sherwood, 2000; Bates and Jackson, 2001). In addition, it could be found that the ratio of improvement of water vapor profile from background is relatively lower than that of temperature profile that is caused by the higher instrument noise of water vapor than CO_2 channels. The enhancement of water vapor profile from the new IASI channels is attributed to the added water vapor channels in the spectral range of 1350 cm^{-1} to 2000 cm^{-1}, since the new IASI channels have more water vapor channels than the operational channels.

In addition to the validation using the retrieved atmospheric profiles, the CSI analysis is also performed in order to understand which the additional channels in new IASI channels contribute the improvement of atmospheric parameters retrieved from 1D-Var. Since the apparent distinction of new IASI channels from

operational IASI channels is the added ozone and IASI band 3 channels, we examined the variation of CSI according to the use of added ozone or IASI band 3 channels. A number of added ozone and IASI band 3 channels in new IASI 200 channels is 11 and 26, respectively. As shown in Figure 21, the CSI of new IASI channels is significantly larger than that of operational IASI channels, but if IASI band 3 channels are excepted the CSI is sharply reduced, meaning that most of the improvement of new IASI channels comes from the added IASI band 3 channels. In contrast, the CSI of new IASI channels except ozone channels (Fig. 21c) is quite similar with that of new IASI channels (Fig. 21b) which means that added ozone channels have a slight improvement of atmosphere parameters from 1D-Var. In addition, the discrepancy between operational and new IASI channels except IASI band 3 channels (Fig. 21a and 21d) could be caused by added water channels.

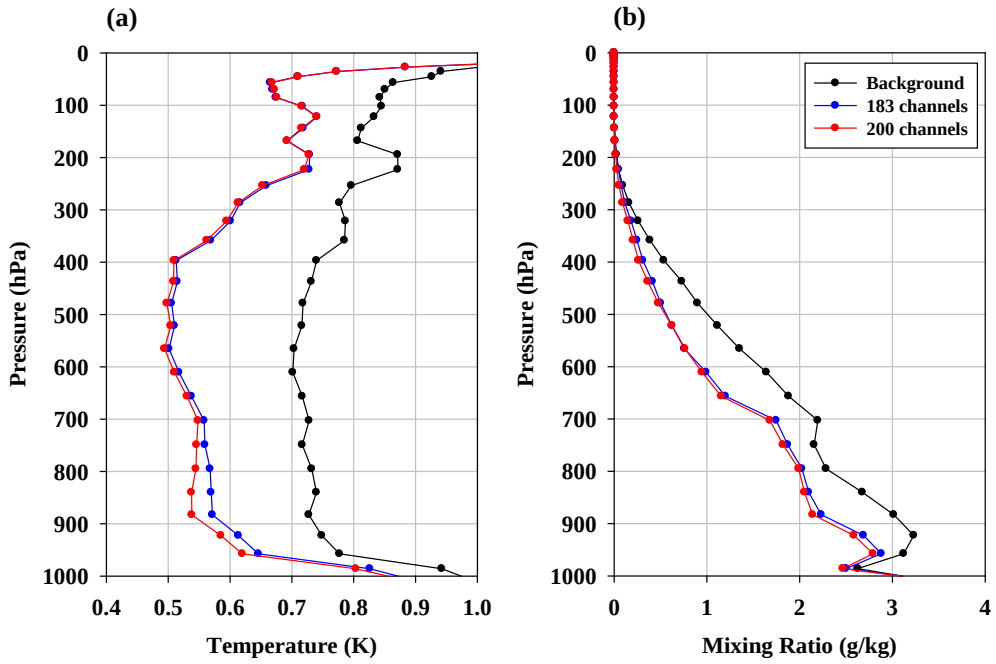

Figure 20. Vertical structure of analysis error (i.e. rmse) of (a) temperature and (b) mixing ratio resulting from the 1 D-Var analysis. The black lines represent the background, blue lines show the case of operational IASI 183 channels, and red lines indicate the case of new selected IASI 200 channels.

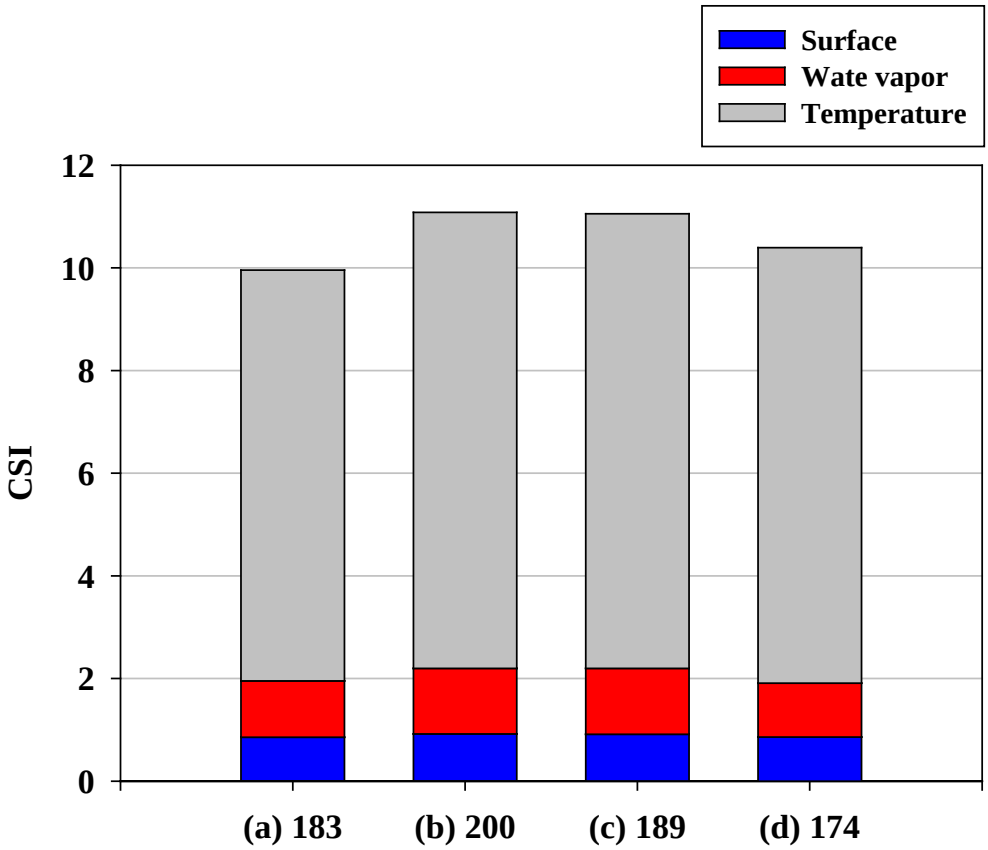

Figure 21. Channel Score Index (CSI) resulting from (a) operational IASI 183 channels, (b) new selected IASI 200 channels, (c) new IASI channels except ozone channels, and (d) new IASI channels except IASI band 3 channels. Color bar indicates the CSIs of a component of temperature (gray), water vapor (red), and surface variables (blue) including skin temperature, surface water vapor, and surface pressure.

4.6. Impact of new IASI channels on the UM global forecasting

Global model assimilation trials were carried out for a period from July 4–August 7, 2013, to examine the impact of the new set of IASI channels on the analysis fields of the operational assimilation system and resultant forecasts. The trials consisted of control and experiment runs. The 183 operational IASI channels were used in the control run and were replaced by the new 200 IASI channels in the experiment run. All the baseline observations were the same for both the runs, which were performed with the operational global UM at the Met Office.

4.6.1. Assessment of first-guess departure

For assimilating the satellite observation data, a quality-control process known as OPS in the operational UM system is performed. Satellite-observed radiances are compared with the radiances simulated with collocated model-generated atmospheric fields (called the background) using a radiative transfer model (e.g., RTTOV). Background fields are the UM 6-hour forecasts issued from the previous data assimilation cycle. Through the examination of the departure of simulated radiances from satellite observations, the performance of the background fields can be monitored.

The experiments with the new 200 IASI channels had a substantially smaller

data volume ratio of IASI observations passing through the OPS than the control run. This significantly reduced volume of IASI observations was due to ozone channels included in the new 200 IASI channels. The use of climatological ozone concentration in the UM for the O_3 channel simulations caused larger departures from the observed radiances, resulting in a convergence failure in the 1D-Var process (e.g., Saunders et al., 2013). Thus, O_3 channels are excluded from the 200 IASI channels in the experiment runs, in order to increase the convergence ratio for 1D-Var in the OPS. This results in a selection of 189 channels, which is close to the operationally used 183 channels.

Similar problems may arise from the use of shortwave infrared channels over band 3 during the daytime because of solar contamination. Thus, the nighttime band 3 measurements are used, which are defined as the ones with a solar zenith angle greater than 95°.

In order to examine the impact of new IASI channels on the background field, atmospheric profiles of the background from the control and experiment runs are used as inputs to RTTOV to simulate channel radiances of the Advanced TIROS Operational Vertical Sounder (ATOVS). The ATOVS consists of three instruments: the High Resolution Infrared Radiation Sounder (HIRS), the Advance Microwave Sounding Unit-A (AMSU-A), and the Microwave Humidity Sounder (MHS). The differences in first-guess departure between the control and experiment runs for

the ATOVS channels in four satellite platforms (i.e., MetOp-A, MetOp-B, NOAA-18, and NOAA-19) are shown in Figure 22. The characteristics of ATOVS channels are given in Table 4. From the comparison between the control and experiment runs, little difference is found in the temperature sounding channels of ATOVS (i.e., HIRS channel 2–7, given as channel indices of 2–7 and AMSU-A channel 1–14, given as channel indices of 21–34). By contrast, significant differences are noted in the H_2O channels (HIRS 12 given as channel index of 12, and MHS 3–4 given as 38–39). Considering that these H_2O channels show Jacobian peaks in the upper troposphere, the new channels may provide a bigger impact on the upper-tropospheric humidity field. However, considering the finding that negative biases become larger especially for the upper-tropospheric channels (e.g., channel indices 12 and 38), the experiment run should induce a relatively drier upper troposphere in the 6-hour forecast. The negative biases of the upper-tropospheric humidity channels are persistent throughout the analysis period (Fig. 23). Persistent larger biases are also clear in the experiment run.

Such drier biases shown in the experiment run are also evident from the comparison with simulations for other infrared sensors (i.e., the Atmospheric Infrared Sounder (AIRS) and Cross-track Infrared Sounder (CrIS) hyperspectral sounders) with the same background fields. In the UM forecasting, 324 AIRS and 399 CrIS channels over CO_2, H_2O absorption bands, and window region are

included in the system (Cameron et al., 2005; Gambacorta and Barnet, 2013). However, similar to the IASI channel selection, some of the AIRS and CrIS channels were not used for data assimilation – the unused channels are indicated as empty spaces in Figure 24. The experiment run shows larger negative mean biases for H_2O channels (e.g., AIRS 174–235 and CrIS 185–312 channels), compared to the positive biases shown in the control run. In particular, the larger negative biases are located over higher channel numbers, which represent channels sensitive to the upper-tropospheric humidity. Such relative dry biases for the experiment run are consistent with the ATOVS results shown in Figure 22.

As shown in the first-guess departures of ATOVS, CrIS, and AIRS H_2O channels, the new IASI channel selection causes larger biases, and the larger biases suggest inferior performance in the assimilated water vapor field. However, considering that the biases between the satellite observations and model backgrounds are subtracted from satellite observations before use for data assimilation (called 'bias correction'), the larger H_2O channel biases might be related to such bias correction.

Here, for the bias correction, we use the correction scheme employed by the UM OPS, which is based on the bias correction scheme developed by Harris and Kelly (2001) for satellite measurements. The scheme uses the geometric thicknesses of 850–300 hPa and 300–50 hPa layers in the model atmosphere as

predictors to regress to the departures between IASI observations and background-derived radiances. Once regression coefficients and an intercept point are determined for each of all IASI 314 selected channels, biases for used 189 IASI channels are predicted by inserting two layer thicknesses into the obtained regression equation. Since our main aim is to examine how the new set of channels improves the forecasts through the improvement of background field, recalculation of bias correction coefficients for the current experiment may not be necessary and therefore the same bias correction coefficients are used for both the experiment and control runs.

The bias correction mostly removes biases related to instrument errors and uncertain radiative transfer modeling. These biases are variable with time because of the sensor degradation or upgrade of the radiative transfer model. However, in two trial runs (control and experiment run) in this study, the same satellite data are used along with the same version of radiative transfer model. Because sensor-related errors and systemic errors are same for both two runs, the same bias correction coefficients can be used for both runs.

Here, using radiosonde observations, we examine above mentioned bias correction as a possible cause for the larger negative H_2O channel departures shown in Figures 22 and 24. Mean biases of first-guess departures obtained from radiosonde observations in the assimilation cycle are shown in Figure 25. Negative

biases are clear for both the control and experiment runs, suggesting that backgrounds (i.e., 6-hour forecasts) are humid nearly all over the troposphere, compared to radiosonde observations (Fig. 25). Much larger humidity biases of the control run over the upper troposphere are significantly different from the near-zero bias shown in the first-guess departures (Figs. 22 and 24), and is thought to be due to the bias correction made in the OPS. For the satellite-measured radiance assimilation, biases are removed from the observations in order to make observations consistent with model backgrounds.

Thus, the apparent near-zero biases of the H_2O channels in Figures 22 and 24 do not necessarily mean a nearly unbiased moisture condition in the control run, as noted in the humid condition in Figure 25. The bias correction in the control run should have been a positive brightness temperature subtraction from the satellite measurements (i.e., colder H_2O channel temperatures representing humid condition). The experiment run, on the other hand, indicates a less humid condition in comparison to the control run (Fig. 25). Since the coefficients used in bias correction for both the experiment and control runs are the same, the apparent departures should be negative as in Figures 22 and 24, because of the less humid condition as shown in Figure 25. Thus, we conclude that the larger negative departures shown in the H_2O channels (Figs. 22 and 24) are due to the use of the same bias correction for both the control and experiment runs, and the near-zero

biases in the control run should not mean an unbiased humidity condition.

Table 4. Characteristics of ATOVS channels used in the UM data assimilation.

Instrument	Channel	Channel Index	Wavenumber (cm^{-1})	Wavelength (μm)	Principal absorber
HIRS	2	2	679	14.7	CO_2
HIRS	3	3	691	14.5	CO_2
HIRS	4	4	704	14.2	CO_2
HIRS	5	5	716	14.0	CO_2
HIRS	6	6	732	13.7	CO_2
HIRS	7	7	748	13.4	CO_2
HIRS	8	8	898	11.1	Window
HIRS	10	10	1217	8.3	H_2O
HIRS	11	11	1364	7.3	H_2O
HIRS	12	12	1484	6.7	H_2O
HIRS	13	13	2190	4.57	CO_2
HIRS	14	14	2213	4.52	CO_2
HIRS	15	15	2240	4.46	CO_2
Instrument	**Channel**	**Channel Index**	**Frequency (GHz)**	**Polarization angles**	**Principal absorber**
AMSU-A	1	21	23.8	90–θ	Window
AMSU-A	2	22	31.4	90–θ	Window
AMSU-A	4	24	52.8	90–θ	O_2
AMSU-A	5	25	53.6	θ	O_2

AMSU-A	6	26	54.4	θ	O_2
AMSU-A	7	27	54.9	90–θ	O_2
AMSU-A	8	28	55.5	θ	O_2
AMSU-A	9	29	57.29	θ	O_2
AMSU-A	10	30	57.29 ± 0.217	θ	O_2
AMSU-A	11	31	57.29 ± 0.322 ± 0.048	θ	O_2
AMSU-A	12	32	57.29 ± 0.322 ± 0.022	θ	O_2
AMSU-A	13	33	57.29 ± 0.322 ± 0.010	θ	O_2
AMSU-A	14	34	57.29 ± 0.322 ± 0.0045	θ	O_2
MHS	3	38	183.31 ± 1.0	θ	H_2O
MHS	4	39	183.31 ± 3.0	θ	H_2O
MHS	5	40	190.31	90–θ	H_2O

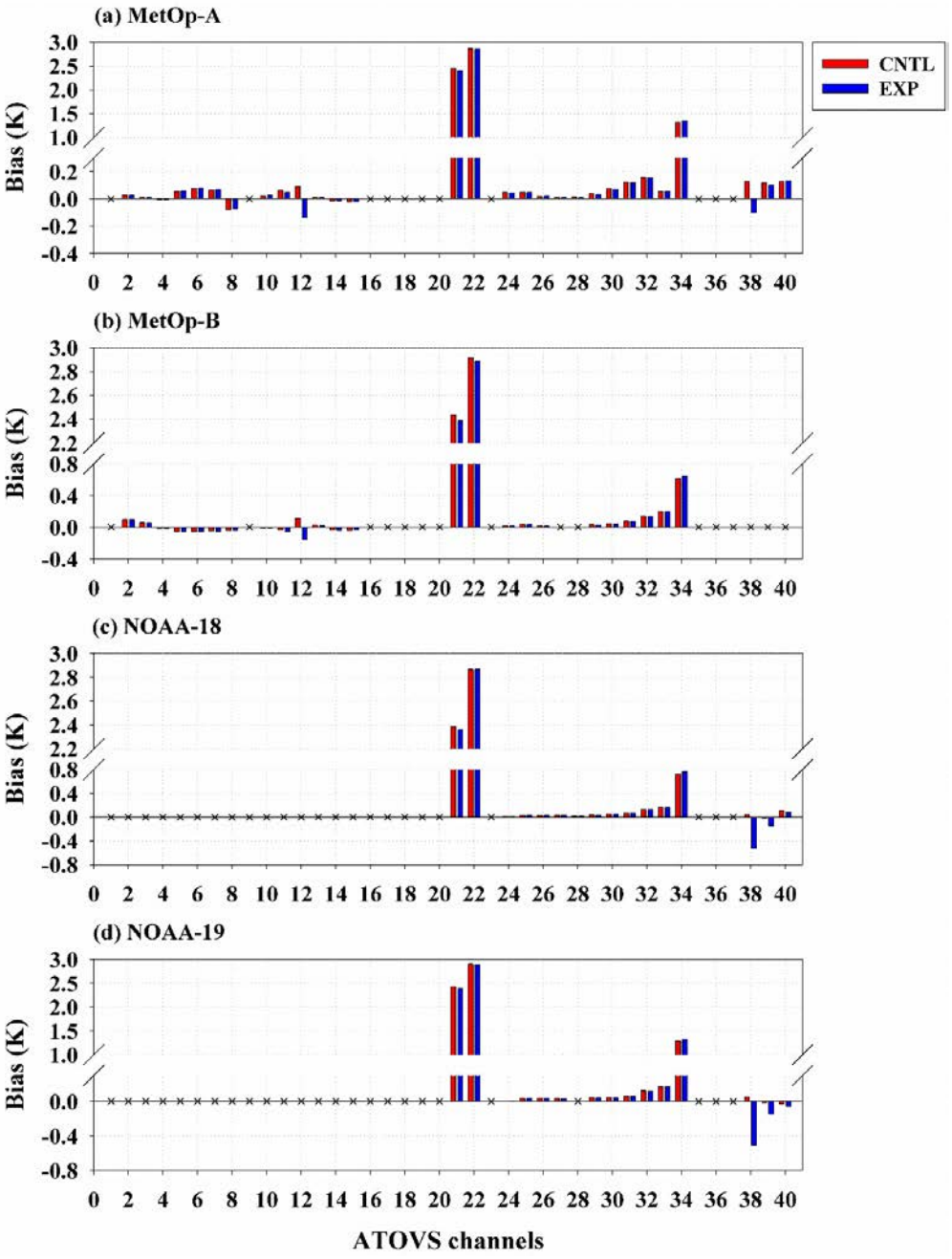

Figure 22. Mean first-guess departures for ATOVS channels onboard (a) MetOp-A, (b) MetOp-B, (c) NOAA-18, and (d) NOAA-19 satellites for a period from July 4–August 7, 2013. The red and blue colors indicate results from the control and experiment runs, respectively. Channel characteristics are shown in Table 4.

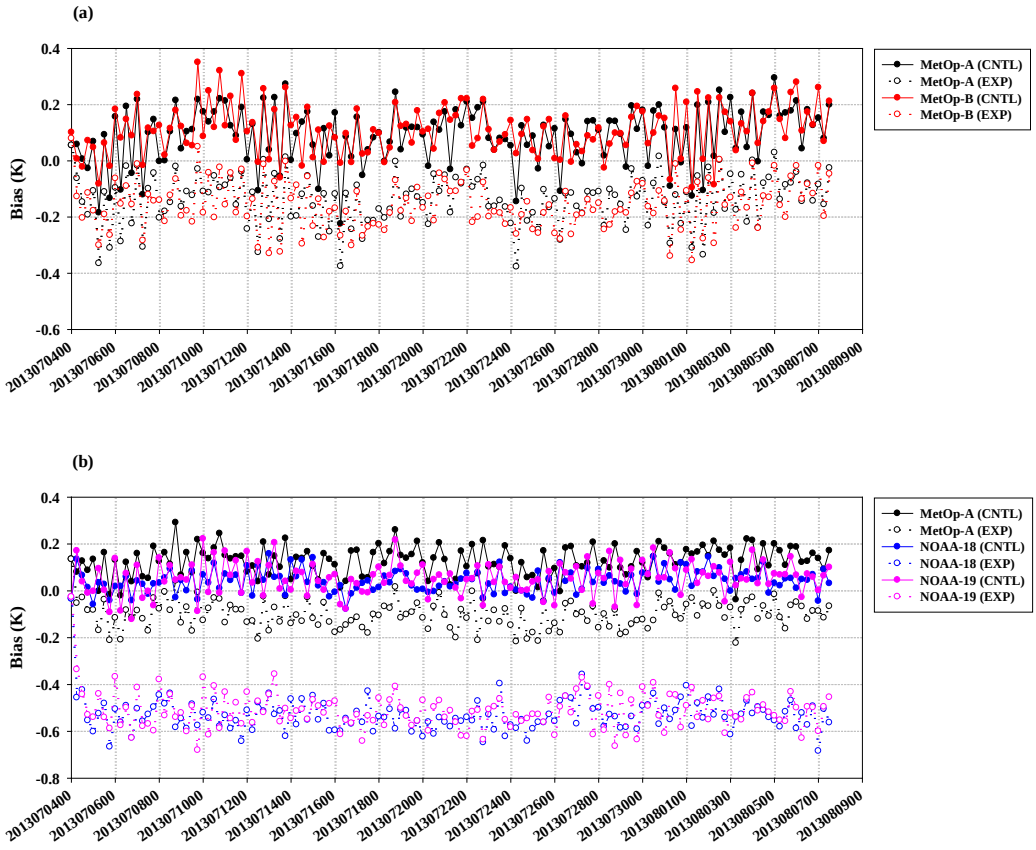

Figure 23. Time series of mean first-guess departures for (a) HIRS channel 12 and (b) AMSU-B channel 3 onboard MetOp-A (black), MetOp-B (red), NOAA-18 (blue), and NOAA-19 (pink) for a period from July 4–August 7, 2013. Solid line (with closed dot) and dashed line (with open dot) indicate the control and experiment runs, respectively.

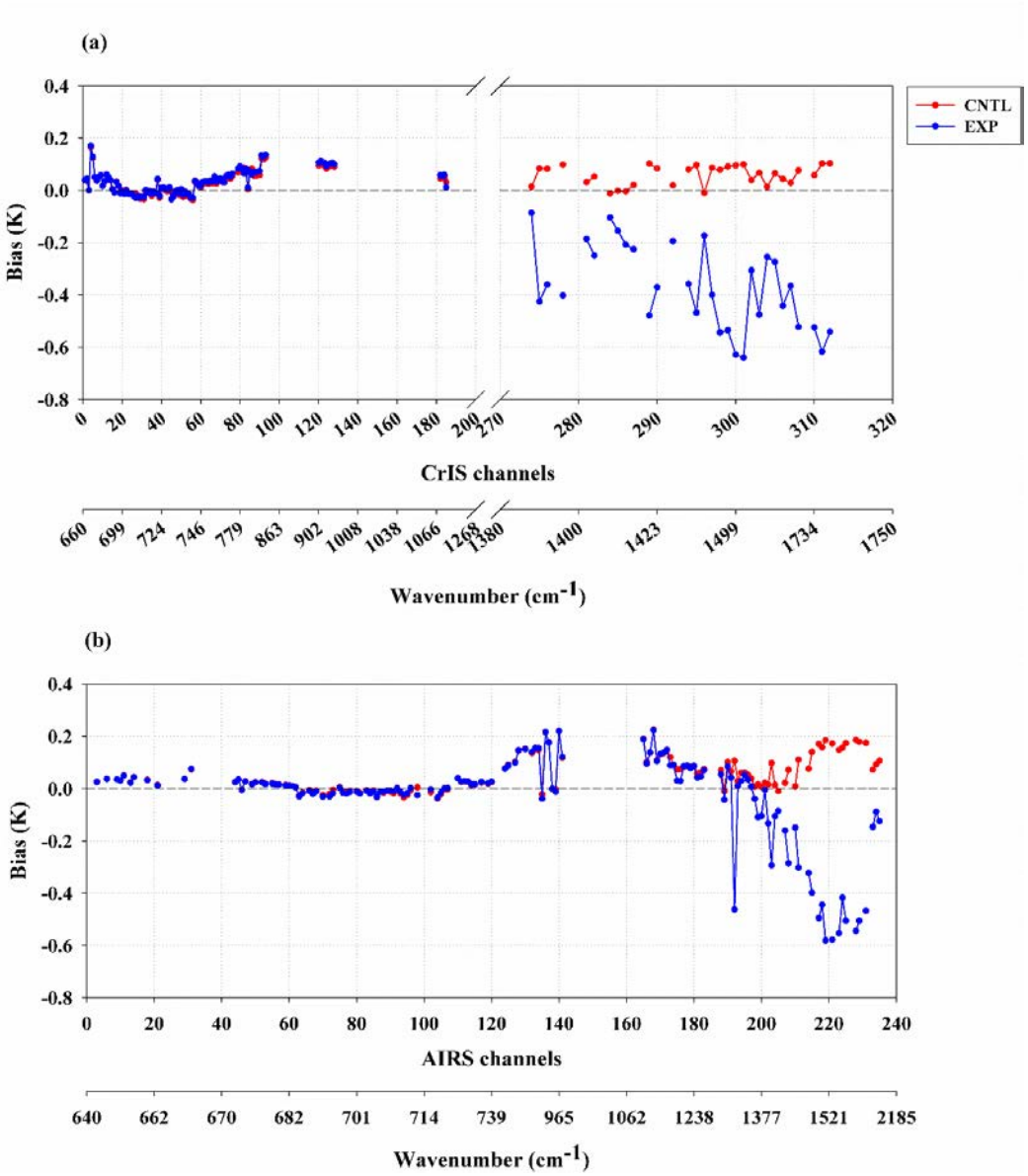

Figure 24. Mean first-guess departures for (a) CrIS and (b) AIRS channels for a period from July 4–August 7, 2013. The red and blue lines indicate the results from the control and experiment runs, respectively.

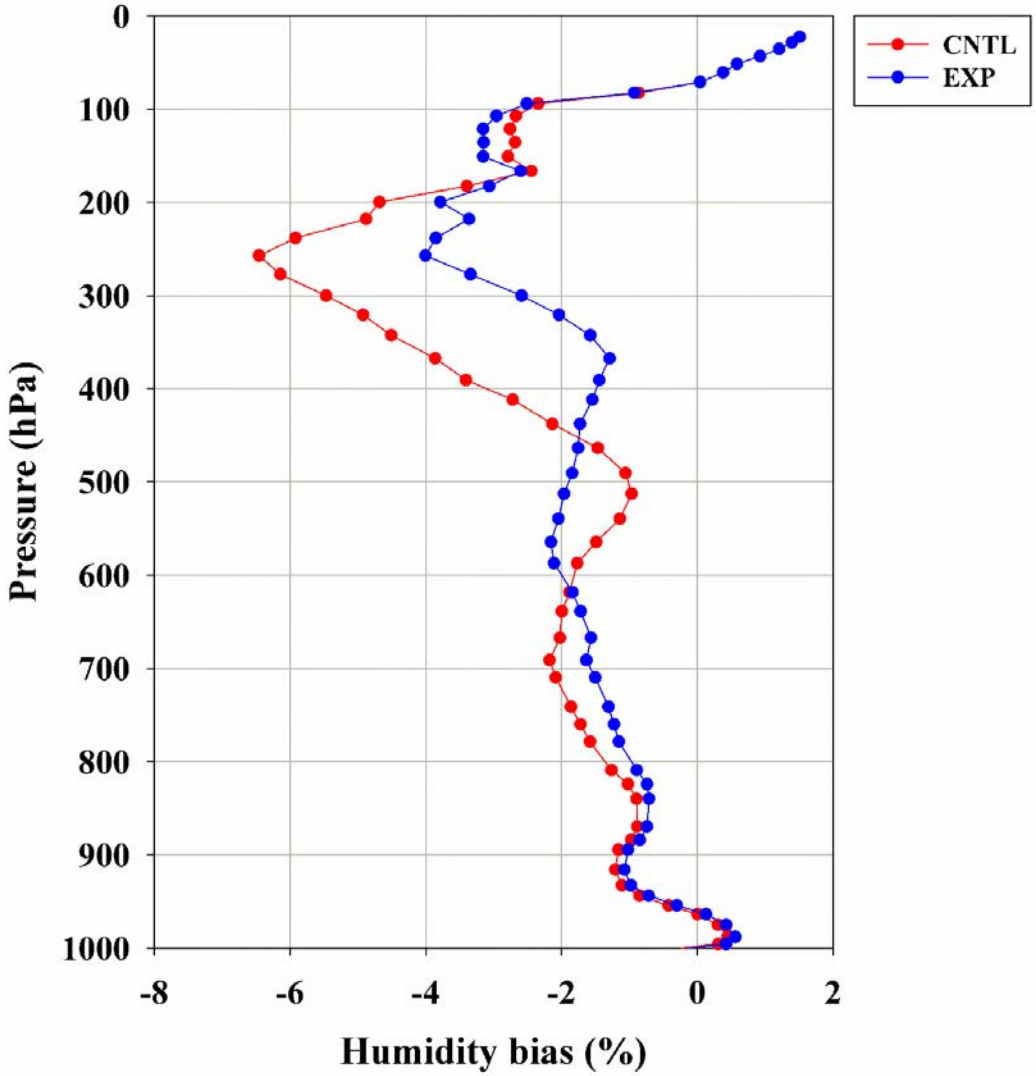

Figure 25. Mean first-guess departures for radiosonde relative humidity profiles for the control (red line) and experiment (blue line) runs.

4.6.2. Impact assessment using global model assimilation trials

In order to examine the impact of the new IASI channels on the UM global forecasts, departures of forecast atmospheric parameters obtained from each trial run against radiosonde observations were calculated. The score called the global NWP index is used to measure the forecast impact in the UKMO UM system; it combines the scores of particular atmospheric parameters at each forecast hour. The total index is the sum of the scores from all the atmospheric parameters. Details regarding how to derive the scores are found in appendix A of Rawlins et al. (2007).

The global NWP index calculated for the experiment, shows that the newly selected IASI channels have an overall neutral impact on the forecast improvement for the selected atmospheric parameters, compared to the control run (not shown). Since the main difference between the new IASI channels and the currently used operational IASI channels is the use of different H_2O channels and band 3 channels, it is not surprising to find little improvement in the NWP index, based on parameters such as mean-sea-level pressure, 500 hPa geopotential, and winds at 250 hPa and 850 hPa.

In contrast, the forecast temperature and relative humidity parameters, which are not part of the NWP index calculation, show a significant improvement especially above 500 hPa. Figures 26 and 27 show the mean biases of temperature

and relative humidity at 500 hPa and 850 hPa levels, in the range of T+0 to T+72 forecast hours. The mean biases of the forecast temperature at 500 hPa level were smaller for the experiment run, compared to those from the control run, in both the Northern and Southern Hemispheres, although the mean bias in the Southern Hemisphere at T+36 forecast hour became larger first. In contrast, the forecast error for the temperature at 850 hPa shows a slight degradation for the experiment run.

In terms of humidity in the control and experiment runs, the positive biases of relative humidity at 250 hPa level are significantly reduced by up to 4% in the experiment run in both the hemispheres, as shown in Figures 27a–b. However, for the mean bias difference at the 850 hPa level for temperature, the difference in relative humidity also shows a less significant impact. The RMSEs of relative humidity are slightly reduced at 250 hPa level in the experiment run but quite similar at 850 hPa level (Fig. 28). It is clear that the improvement in the relative humidity forecasts mainly occurred in the upper troposphere from 500 hPa to 200 hPa. A substantial bias reduction is evident in the analysis at the T+0 forecast hour. By contrast, in the mid-troposphere from 600 hPa to 500 hPa, the mean biases are larger for the experiment run at T+0 and T+24 forecast hours (Fig. 29a and b). The humidity RMSEs are mainly reduced in the upper-troposphere from 350 hPa to 200 hPa at the T+0 and T+24 forecast hour but the RMSE reduction at the T+48

and T+72 forecast hour appears to be insignificant (Fig. 30). As expected, the humidity biases and RMSEs increase with the forecast hour in both the runs, even though the mean biases and RMSEs in the experiment run are smaller than those in the control run.

Considering that the main difference between the new channels and the UKMO 183 operational channels lies in the use of the new IASI H_2O channels and band 3 channels, the improvement in the temperature and relative humidity forecasts in the upper troposphere can be attributed to the use of these new additional channels. Since weighting functions of the selected IASI band 3 channels are mostly located in the lower troposphere, we expect improved temperature forecasts in the lower troposphere. However, there is only a small change in the lower troposphere from the use of the new channels, which suggests that the impact of the band 3 channels may be minimal, probably due to the high noises in these channels. Thus, the overall improved forecasts for temperature and humidity in the upper troposphere are primarily due to the newly selected H_2O channels.

It is also important to note that the difference in the relative humidity bias in the upper troposphere between the control and experiment runs became smaller with forecasting hour, except around the 250 hPa level at T+72 forecast hour (Fig. 29). The assimilation of additional channels in the experiment run produces a

humidity analysis closer to the radiosonde values. But such bias reduction disappears quickly and the humidity bias tends to return to the level of the control run.

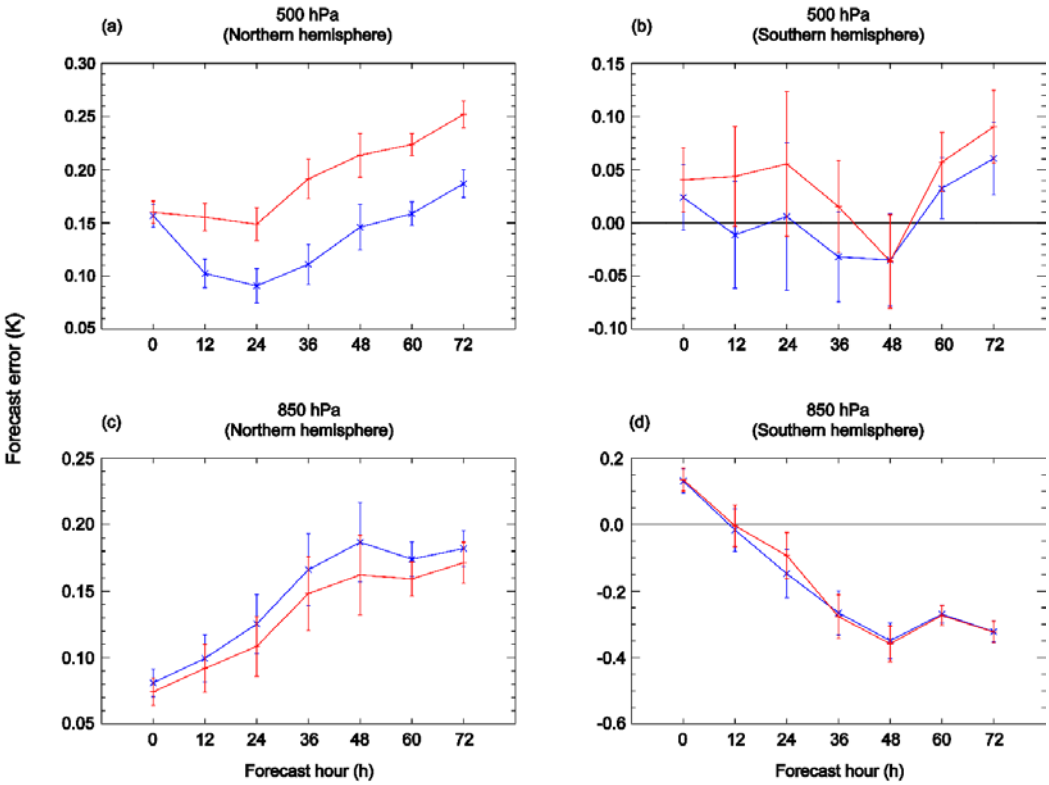

Figure 26. Mean biases of 500 hPa and 850 hPa temperature forecast errors between forecasts and radiosonde observations over the Northern Hemisphere (left panels) and Southern Hemisphere (right panels). Red and blue indicate the control and experiment runs, respectively. Vertical bars representing one standard deviation of the uncertainty are overlaid.

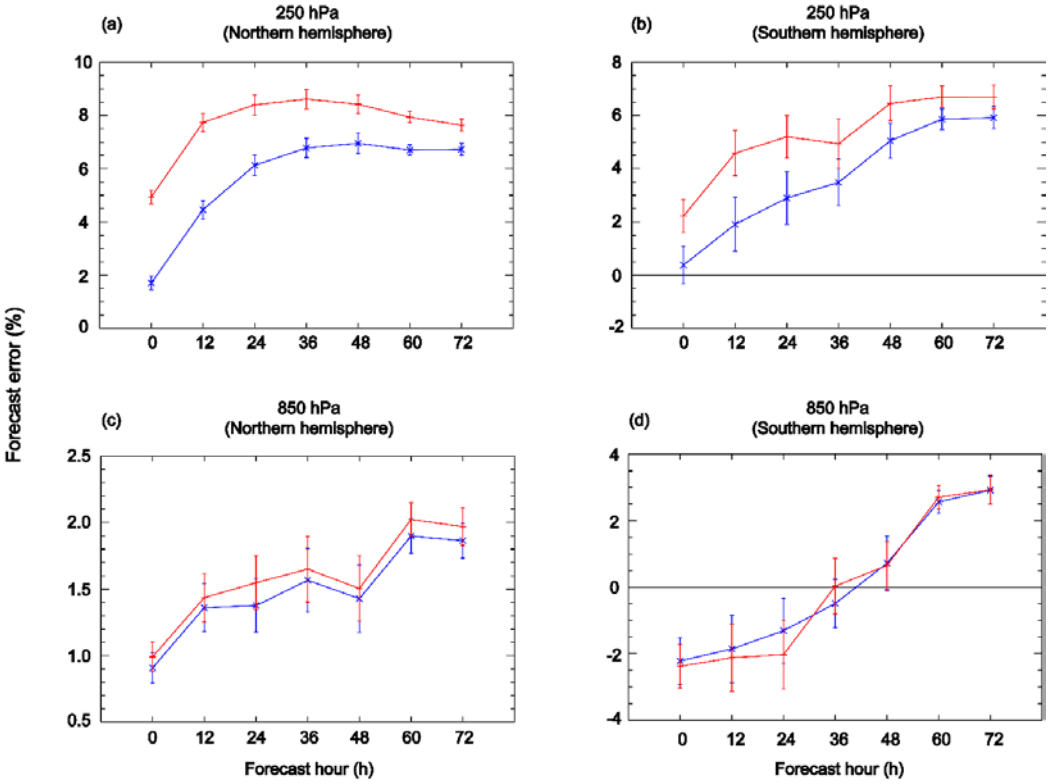

Figure 27. Same as in Figure 26 except for relative humidity forecast errors.

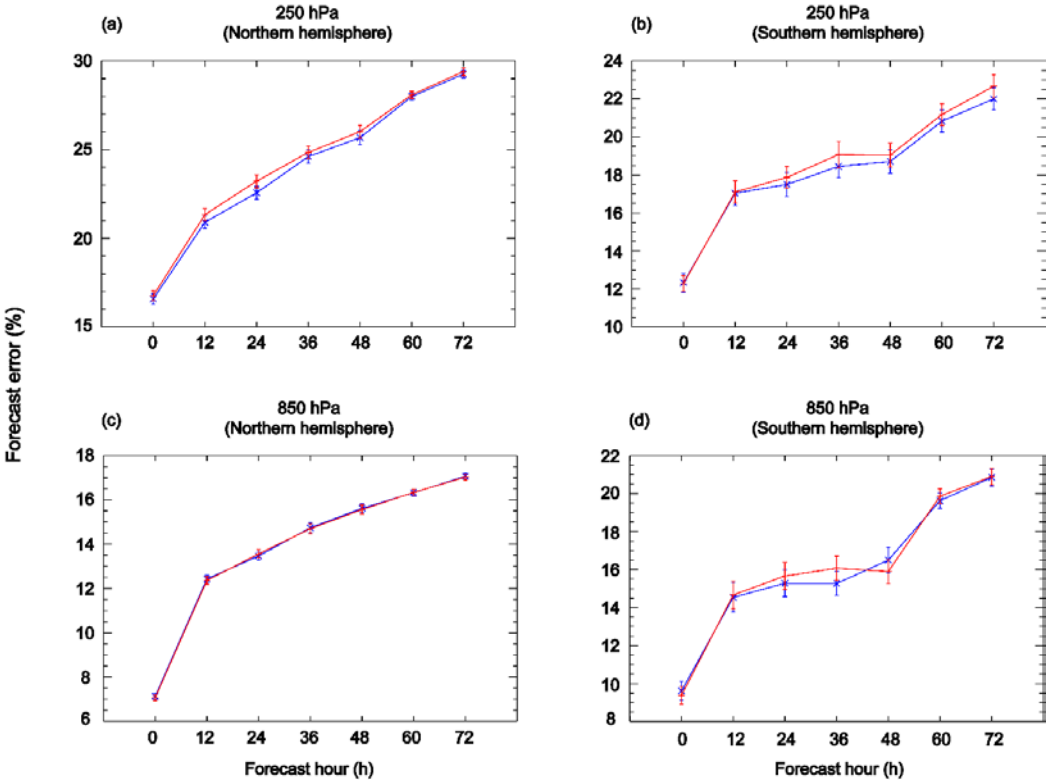

Figure 28. The root-mean-squared error (RMSE) of 250 hPa and 850 hPa relative humidity forecast errors between forecasts and radiosonde observations over the Northern Hemisphere (left panels) and Southern Hemisphere (right panels). Red and blue indicate the control and experiment runs, respectively. Vertical bars representing one standard deviation of the uncertainty are overlaid.

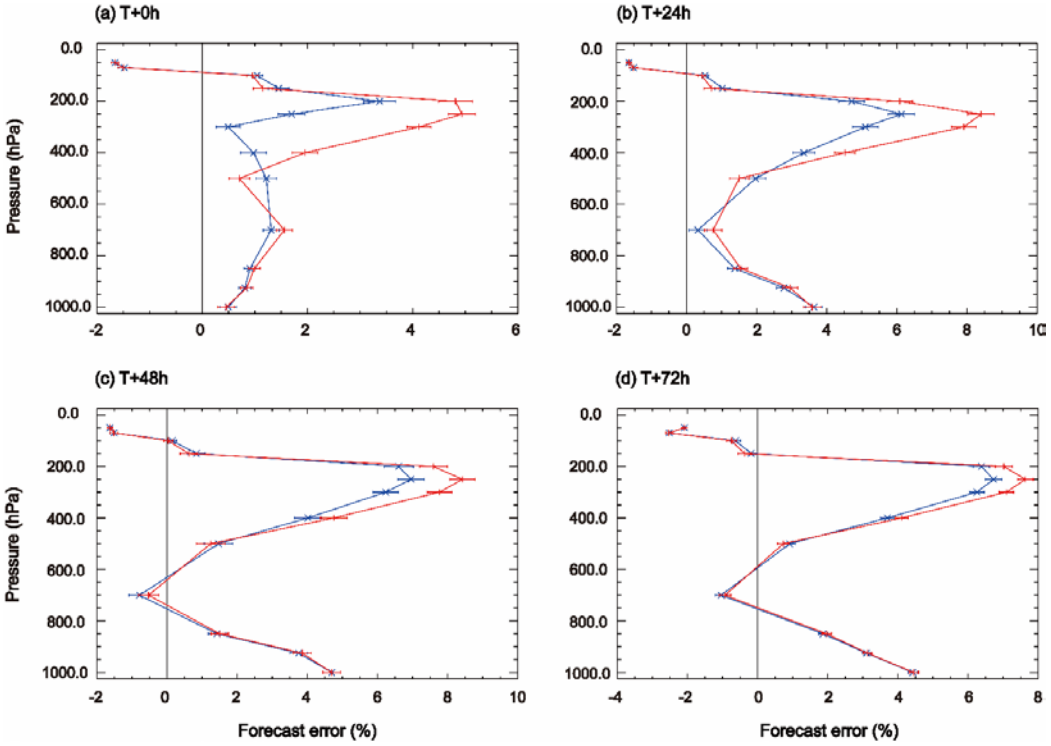

Figure 29. The vertical structure of mean bias of the relative humidity forecast errors between forecasts and radiosonde observations at the T+0 to T+72 forecast hours. Red and blue indicate the control and experiment runs, respectively. Horizontal bars indicate one standard deviation of the uncertainty.

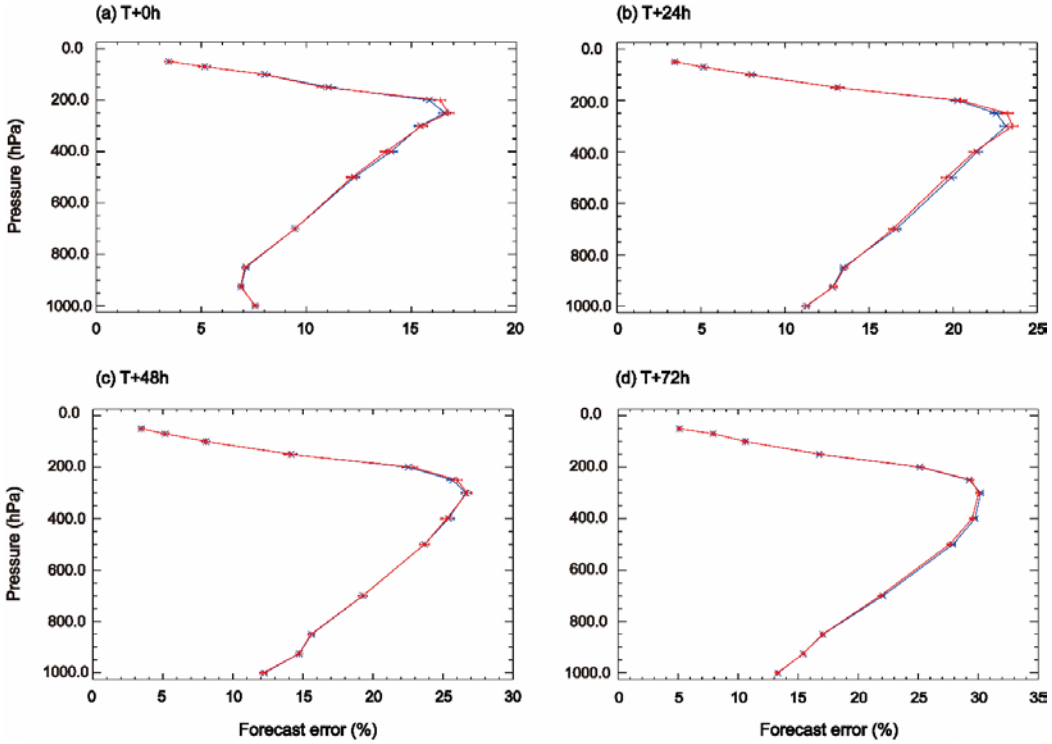

Figure 30. Same as in Figure 29 except for the root-mean-squared error (RMSE) of the relative humidity errors.

4.7. Impact of new IASI channels on the UM precipitation forecast over East Asia

In order to examine the impact of newly selected IASI channels on the UM precipitation forecasts for the rainfall events over East Asia, global assimilation trials were conducted for a period from June 15–July 31, 2015, using the global data assimilation system at KMA. The trials were comprised of control and experiment runs. Operational 183 IASI channels were used for the control run. And in the experiment run those channels were replaced by the new 200 IASI channels. Same as in the trials for examining the impact of new IASI channels on the UM global forecasts, ozone channels in the new IASI channels were excluded and the shortwave infrared channels over the IASI band 3 were used only at night time. All the same baseline observations including ground-based and satellite observations were used for the both runs.

4.7.1. Impact on the precipitation forecast over East Asia

Even though the trials were carried out using the global data assimilation system, we focused on evaluating the precipitation forecast for rainfall events that occurred around the Korean peninsula (32.5°N–40°N, 120°E–135°E). In addition, in order to avoid considering horizontally small-scale rain events, the rain events whose rain area is larger than about 50000 km^2 were chosen for the analysis, using

the precipitation product retrieved from satellite measurements.

Used satellite-based precipitation products were obtained from the global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). The GPM mission is the international satellite mission initiated by the National Aeronautic and Space Administration (NASA) and the Japan Aerospace and Exploration Agency (JAXA). The GPM spacecraft has 65° inclination angle and 407 km altitude orbit, which provides a broad latitudinal spatial coverage. Because the GPM is not locked into a sun-synchronous polar orbit, it provides frequent crossings with other constellation satellites, and facilitates diurnal sampling of precipitation. For the GPM, two observation instruments are onboard the core platform. One is a Ka/Ku-band Dual-Frequency Precipitation Radar (DPR) developed by JAXA, and the other is the GPM Microwave Imager (GMI) developed by the NASA (Draper et al., 2015).

The IMERG algorithm is intended to inter-calibrate, merge, and interpolate various satellite microwave precipitation estimates, together with microwave-calibrated infrared (IR) satellite estimates, precipitation gauge analyses. IMERG is designed to compensate for the limited sampling data available from single low-Earth-orbit (LEO) satellites by using as many LEO satellites as possible, and then filling in gaps with geosynchronous-Earth-orbit (GEO) IR estimates. There are three ways to produce the precipitation products in IMERG algorithm. First, the

LEO-passive microwave (LEO-PMW) data are morphed (linear interpolation following the GEO-IR-based feature motion). Second, GEO-IR precipitation estimates are included using a Kalman filter when the LEO-PMW are too sparse. Finally, precipitation gauge analyses are used to provide crucial regionalization and bias correction to the satellite estimates (Huffman et al., 2015). Various types of precipitation products are specified to satisfy the demand of users. In this study, we used the precipitation product at a temporal resolution of 30 min and spatial resolution of 0.1°, which is "Late" multi-satellite product ~12 hours after observations time.

The accuracy of precipitation forecast for two trial runs was evaluated by use of the threat score (TS) and "Bias" as an evaluation index, which are generally used for dichotomous variables (e.g., precipitation). We defined the East Asia region (25°N–50°N, 115°E–140°E) as an analysis domain. The TS and bias can be calculated using either observations (i.e., IMERG precipitation product) or model analysis produced for each trial run as a reference.

$$TS = \frac{Hits}{Hits + Misses + False\ Alarms} \tag{13}$$

$$Bias = \frac{Hits + False\ Alarms}{Hits + Misses} \tag{14}$$

"*Hits*" represents the intersection area of precipitation between the model forecast and reference area, "*Misses*" is the precipitation area in the reference data, which was missed in the forecast. "*False Alarms*" indicates the precipitation area in the forecast, which was not assigned as the precipitation area in reference data. Thus, the TS has a value between 0 and 1, and a higher (lower) value represents better (poorer) forecast performance for each trial run. A positive "Bias" means that the forecasted precipitation area is wider than the observed area, and vice versa.

For the verification results against observations (observations verification), the mean TS values for the experiment run are larger than that for the control run in the range of T+0 to T+48 forecast hours except T+12 forecast hour (Fig. 31). However, for both runs, the mean TS values tend to decrease gradually with the forecast hour, suggesting that the longer the forecast time, the less accurate the precipitation performance. Similar to the observation verification, the mean TS values, which were calculated using the model analysis (analysis verification), are larger at the experiment run, and for both runs decrease with the forecast time. Compared with the observation verification, however, the magnitude of TS value at the analysis verification is consistently larger at T+12 to T+48 forecast hours. It seems to be induced by the dependency between the analysis and the forecast, because the analysis at each trial run is employed to evaluate its own forecast field

in which forecast time is consistent with the analysis time.

Figure 32 shows the bias of forecasts from two trial runs against observations. The biases for the control run are about 1.10 in the period of T+0 to T+36 forecast hours, and at T+48 hour biases slightly increase. However, for the experiment run, the biases decrease up to 1.07 except for the T+48 forecast hour. The bias reduction suggests that the tendency to overestimate the precipitation area in the model is partially alleviated by using new IASI channels in the assimilation system.

Considering that the main difference between the new IASI channels and operational 183 channels is the use of new IASI H_2O channels and the channels over IASI band 3, the improvement of precipitation forecasts over East Asia can be attributed to the use of those new channels added in the new selection. Among these new added channels, improved water vapor fields by the use of new H_2O channels appear to play an important role in the forecast improvement, because the water vapor is the main initial element of causing precipitation.

In order to evaluate the water vapor field at the initial condition (analysis), we calculated the total precipitable water (TPW) between the surface and model top pressure (0.01 hPa) at the initial conditions for two trial runs, and then verified the TPW using satellite-based TPW products. Used satellite-based TPW data were obtained from the microwave integrated retrieval system (MiRS) managed by the National Atmospheric and Oceanic Administration (NOAA). MiRS algorithm is

based on the 1D-Var inversion scheme with the Community Radiative Transfer Model (CRTM) as the forward and adjoint operators (Boukabara et al., 2010). The first guess (background) is generated throughout a multi-linear regression algorithm, developed by collocating satellite measurements with ECMWF analyses. The algorithm is applied to the microwave measurements from the AMSU and MHS sensors onboard NOAA-18, NOAA-19, MetOp-A, MetOp-B platform, and the Special Sensor Microwave Imager/Sounder (SSMI/S) from DMSP-F16 platform. Although the TPW is retrieved over both land and ocean, we only used the TPW retrieved over ocean because the retrieval error over land is large due to the emissivity uncertainty of the land surface.

The time series of the mean TPW bias of two trial runs against the satellite-based TPW are shown in Figure 33. Although the positive mean biases are consistently represented at both runs for the trial period (July 1–July 31, 2015), the biases are significantly reduced at the experiment run. In particular, the bias reduction is represented evidently in the Northern Hemisphere. In addition, in order to investigate the altitude at which the bias reduction occurs, we compared the mixing ratio between two trial runs. Results indicate that the bias reduction is mainly due to the decreased amount of water vapor in the lower troposphere at the experiment run. Because, among new added channels, some lower-peaking H_2O channels are positioned over IASI band 3, the decreased bias seems to be attributed

to the use of these H_2O channels.

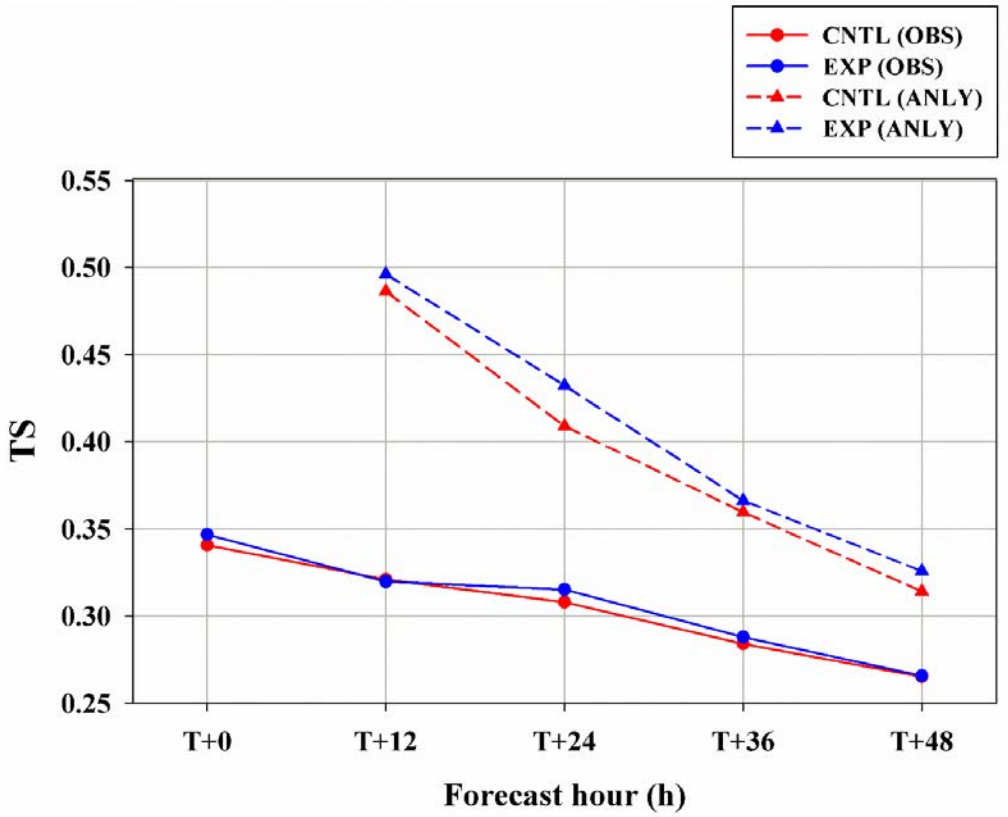

Figure 31. The mean threat score (TS) of precipitation forecast at the T+0 to T+48 forecast hours for a period from June 16–July 31, 2015. Solid and dashed lines indicate the verification results using observations (e.g., satellite-based precipitation product) and model analysis at each run, respectively. Red and blue represent the control and experiment runs.

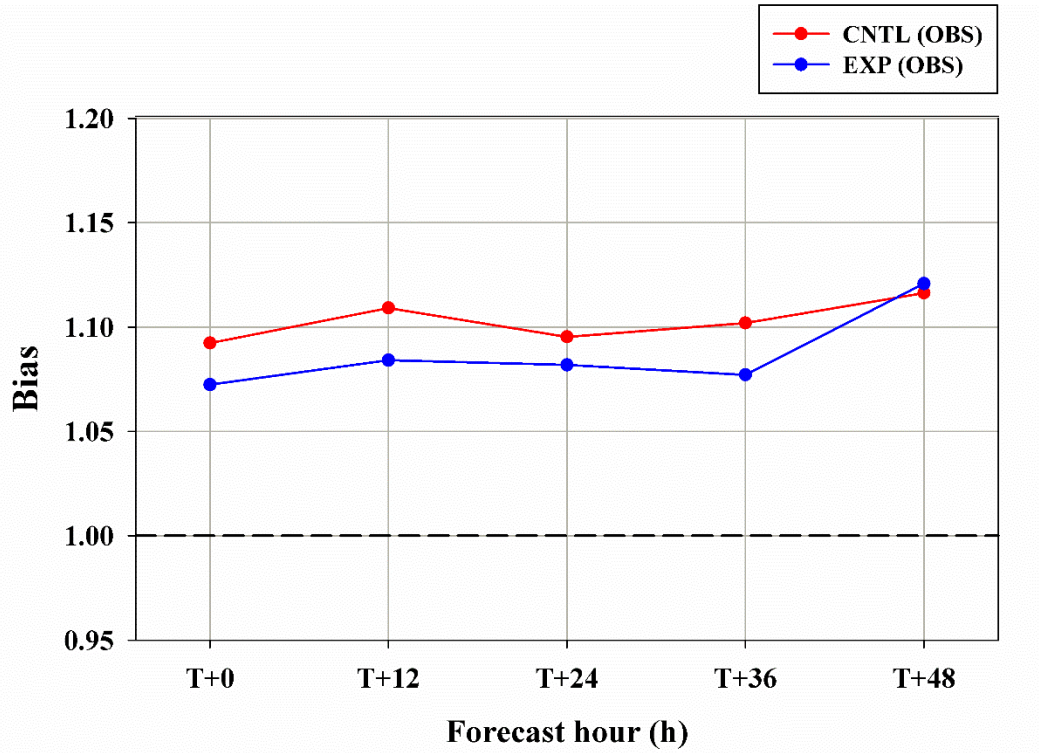

Figure 32. The “Bias” of precipitation forecast at the T+0 to T+48 forecast hours for a period from June 16–July 31, 2015. Red and blue represent the control and experiment runs.

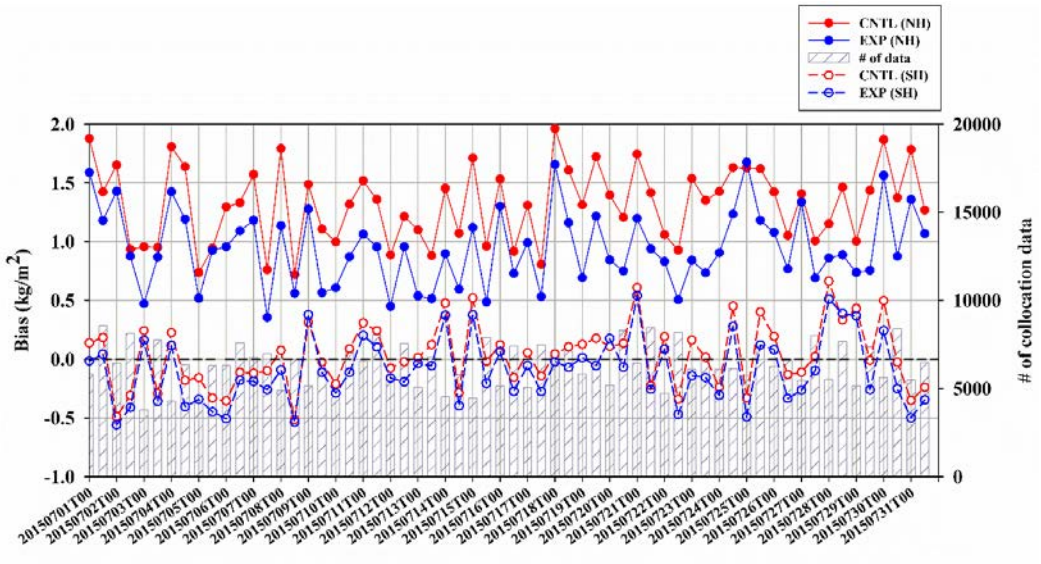

Figure 33. Time series of mean bias of total precipitable water (TPW) between the model and the satellite-based product for a period from July 1–July 31, 2015. Solid and dashed lines indicate the bias in the Northern Hemisphere and the Southern Hemisphere, respectively. Red and blue represent the control and experiment runs. Bars means the number of collocated data.

4.7.2. Impact on the precipitation forecast depending on the rain type over East Asia

In Song and Sohn (2015), heavy rainfall events over the Korean peninsula during the summers of 2004–2011 were classified into two rain types (i.e., warm-type and cold-type rainfall) with the *K*-means clustering algorithm. The radar reflectivity profiles observed by the Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (PR) were employed in the *K*-mean clustering algorithm. Song and Sohn (2015) further demonstrated that the cold-type heavy rainfall seems to result from high cloud system developed via convective instability, which include abundant ice crystals in the upper part of cloud. In contrast, warm-type heavy rainfall appears to be induced by the medium-level cloud system, in combination with strong moisture convergence derived by the horizontal wind field in the humid atmospheric environment over the Korean peninsula. Thus, one of main discrepancies between warm-type and cold-type heavy rainfall may be the structure of cloud system producing certain type of rainfall.

Since the classification of rain type was performed for the heavy rainfall cases during the summers of 2004-2011 in Song and Sohn (2015), it may be ideal to repeat the classification process to divide the rainfall events during the summer of 2015 into two rain types, which is the period of the data assimilation trial experiment. However, because the TRMM measurements were terminated on

April 2015, the TRMM products are not available after the summer of 2015. In addition, even though the GPM continuously provides the radar reflectivity from the DPR sensor, the *K*-mean clustering algorithm may not be applied because of a small set of GPM data. In this study, therefore, we analyzed the results of trial experiments, based on the classification results performed by using the surface rain gauge measurements and lightening measurements (J. Ryu, Personal communication).

Considering that the cold-type rainfall tends to be induced with the high-convective cloud system but the warm-type rainfall occurs with relatively low cloud system, the frequency of lightning accompanied by rainfall can be used as an indicator for classifying the rainfall events into two rain types. It is generally known that the lightening seems to be associated with the development of ice crystal in the upper layer of cloud system. And a large portion of the convective cloud (precipitation) system tends to be produced by ice-based process and thus the positive correlation between convective cloud system and lightning is represented (Peterson and Rutledge, 1998; Seity et al., 2001; Soriano at al., 2001). Therefore, it may be reasonable to classify the rain types based on the lightning frequency.

Two types of measurements taken during the summers of 2002–2015 were used for the classification of rain type. One is the surface rain gauge data from the

Automated Weather Station (AWS) stations, which are densely spread over South Korea. The 1-hourly rain rate (mm h^{-1}) was derived from accumulated AWS rainfall measurements. The other is the cloud-to-ground lightening data, which are measured by the Improved Accuracy from Combined Technology Electronic Stability Program (IMPACT ESP) sensor. The IMPACT ESP data include time and location (i.e., latitude and longitude) of each lightning. After collocating the lightning data with 1-hourly rain rate data at each AWS station, the collocated datasets were made into the grid data of 0.1°x0.1°. There are three steps for the classification of rain type. First, in order to choose the heavy rainfall cases, the case for which the number of heavy rainfall grids (rain rate >10 mm h^{-1}) is greater than 3 is chosen. And then, if the frequency of collocated lightning at each grid is larger (smaller) than the mean lightning frequency at the grid's rain rate, the grid is defined as the cold-type (warm-type) rainfall grid. The mean lightning frequency was derived by averaging the lightning frequency depending on the collocated rain rate during the whole analysis period (summers of 2002–2015). Finally, the rainfall case where the ratio of the number of cold-type grids over total number of heavy rainfall grids is greater (smaller) than 20% is referred as a cold type (warm-type) rainfall case.

Since the UM system used for the trial experiment was designed to produce the long-range forecast at 0000UTC and 1200UTC for a trial period (June 15–July

31, 2015), the number of classified rainfall cases matched with the model forecast becomes small. In particular, the matched cold-type rainfall cases were rare due to the short duration time of the rainfall. As a result, the number of matched warm-type and cold-heavy rainfall cases are nine and four for the trial period (June 15–July 31, 2015), respectively. Same as the verification approach in Section 4.7.1, the TS was used to evaluate the forecast accuracy of precipitation for two rain types.

Figure 34 represents the mean TS values in the range of T+0 to T+48 for the warm-type and cold-type rainfall cases. In terms of the warm-type cases in the control and experiment runs, the TS values of analysis verification in the experiment run are larger than those in the control run regardless of forecast hour, but the TS of both runs tends to decrease with forecast hour (Fig. 34a). Also, for the observations verification, the TS values from the experiment run are slightly larger than those from the control run except for the T+12 forecast hour. Those results suggest that newly selected IASI channels contribute to improved forecast accuracy of warm-type rainfall. By contrast, in the case of cold-type rainfall, newly selected IASI channels have a neutral impact on the precipitation forecast (Fig. 34b). Compared with the warm-type rainfall, the mean TS values decrease overall for both the analysis and observation verification. The tendency of decreasing TS may be due to the difference of spatial scale of rainfall and representativeness error

associated with the spatial resolution of the UM system used for the trial experiment. While the warm-type rainfall tends to cover the wide area with low-altitude cloud system, the spatial coverage of cold-type rainfall was relatively small, suggesting that the TS values of cold-type rainfall are more variable. In addition, because the horizontal resolution of the UM system was about 40 km, the UM system might have limitation to represent the cold-type rainfall with small spatial scale.

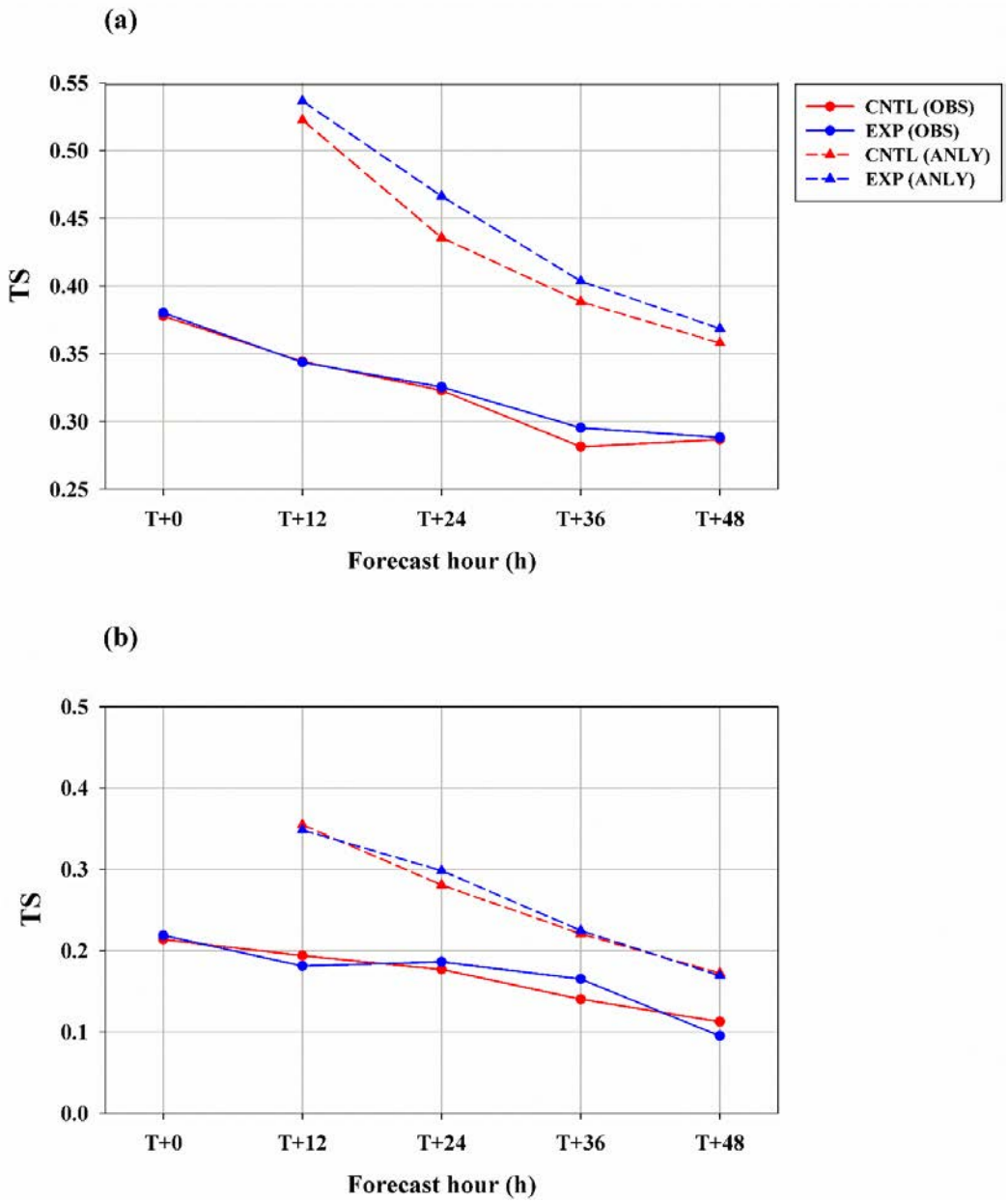

Figure 34. The mean threat score (TS) of precipitation forecast at the T+0 to T+48 forecast hours at (a) warm-type rainfall cases and (b) cold-type rainfall cases for a period from June 16–July 31, 2015. Solid and dashed lines indicate the verification results using observations (e.g., satellite-based precipitation product) and model analysis at each run, respectively. Red and blue represent the control and experiment runs.

5. Summary and Conclusions

At operational weather centers, data assimilation systems, using a diverse range of observations, are an integral part of the state-of-the-art forecasting system. It is, however, important to examine and validate the analysis outputs against in situ measurements for better understanding of the model or assimilation scheme. Here we examine analysis data produced at three operational centers (UKMO, KMA, and ECMWF) where different models and different quality controls for the assimilations are employed. We aimed at examining the overall accuracy of the analysis produced at three centers using common GRUAN-measured temperature and humidity profiles for a one-year period of 2012. Also examined was the impact of the UM version change at the KMA in May 2012 on the performance change in the analysis at the KMA, using the same GRUAN data.

From the comparison of the temperature and water vapor profiles produced by the UM-KMA, UM-UKMO, and ERA-I against collocated GRUAN radiosonde observations, it was revealed that the temperature seemed to agree better between the radiosonde observations and the model outputs at all three centers than the humidity, especially for the upper troposphere.

In the mean error statistical analysis, the UM-KMA and UM-UKMO are found to be comparable, although the RMSDs of the relative humidity were larger for the UM-KMA. In addition, the UM-KMA showed a dry bias for upper levels

in early 2012 (before June), which was significantly reduced after introducing the new UM version in May–June. In contrast, the ERA-I showed the largest positive mean bias for the relative humidity, in particular for the upper layer, while the RMSDs of both temperature and relative humidity were larger than for the UM models throughout the troposphere. The larger RMSDs of the ERA-I were likely due to the larger assimilation time window, the older version of the ECMWF model by the time the ERA-I data were produced, and the residual scale mismatch caused by its coarser spatial resolution in ERA-I relative to the operational NWP models.The humidity errors seemed to be related to the overall humidity, since the RMSDs were proportional to the TPW, being larger for more humid conditions. The similarity between the RMSDs of the analyses and the natural variability of the moisture profile suggested that the NWP systems studied here were less skillful in accurately analyzing humidity for more humid and variable conditions.

We have made a new IASI channel selection with an aim of improving the performance of the UM model forecasts. The method employs a 1D-Var approach with a new defined figure of merit called the CSI. In this procedure, in order to alleviate the known comparatively large influence of water vapor on the channel selection, weight obtained by normalizing each layer's mean water vapor with the mean water vapor at the lowest model layer, is given to each layer's water vapor. Those weights were only given to the diagonal component of the error covariance

matrix because there is no way to give such weights to the cross-covariance components. Also examined is how the CSI results can be compared with results from the ER method when the 1D-Var approach is used with a different figure of merit, ER. The ER method adopted in this study is different from Collard (2007) sequential/linear method because of the non-linear aspect of 1D-Var. However, different from the CSI method, water vapor weights were not given to the error covariance matrix because ER method requires the full error-covariance matrix.

By applying the CSI method, 200 channels were selected out of the 314 provisional IASI channels. Compared with the operationally used 183 channels at the Met Office for the UM data assimilation system, the new channels share 149 channels while other 51 channels are new. Those new channels include H_2O channels representing the water vapor absorption in the middle to upper troposphere, channels in the O_3 absorption band, and channels over the shortwave infrared IASI band 3.

Compared to CSI results, the ER method with 1D-Var selected less channels over the CO_2 absorption band but more channels over the H_2O absorption band. In particular, more upper-tropospheric H_2O channels over 1350–1600 cm^{-1} were selected because contributions of upper tropospheric H_2O channels to the ER are much larger than for channels at any other spectral bands. It seems that the ER method should be applied separately to each important absorption band and put

results together subjectively to give the total number of selected channels. Thus, the CSI method is found to be more objective.

In this study, we tried to select optimal channels from 314 provisional IASI channels that have been operationally received via EUMETCAST in the UM assimilation system. It would have been much desirable if the selection was made from the total 8461 IASI channels using the proposed CSI method. However, the computational burden caused by1D-Var runs appears formidable; even in the case of selecting 200 channels from 314 pre-selected channels, the total number of 1D-Var runs required is 25,740,000 (= (314+313+312+ ··· +115) × 600 profiles).

Being aware of the limitation brought by starting with a small number of pre-selected channels, it is of interest to examine how the channel selection by the CSI method is influenced by a different set of pre-selected channels. It is of further interest to examine how the new set of channels, if there is any, is comparatively informative against the old set (here, for example, EUMETSAT 314 channels vs. 500 channels). As a case of point, we give an effort of selecting new channels from pre-selected EUMETSAT 500 channels (distributed from 15 October 2014), by using the same procedures provided in this paper.

Compared with 200 channels selected from 314 channels (referred to as Collard 200), EUMETSAT 200 channels include more split window channels and low-tropospheric water vapor channels from Martinet's 134 channels and some

additional channels from 43 CO_2 channels of Collard and McNally (2009). In terms of total cumulative CSI, EUMETSAT 200 channels have the CSI 6.7% larger than from Collard 200 channels, meaning that EUMETSAT 200 channels are more informative than Collard 200 channels. It suggests that it may be necessary to select a new subset of channels if the environment is changed, in association with NWP model or supply of satellite data, in order to perform more efficient data assimilation. Considering that all channels across the absorption bands should not be equal, determining most important channels in sequence must be a valuable effort.

To assess the impact of the new 189 channels on the data assimilation system compared to the operationally used channels, an experiment was conducted for a summer period from July 4–August 7, 2013, using the UKMO global UM system. O_3 channels were excluded in this experiment because the UM model does not explicitly forecast the ozone concentrations. While the control run used the UKMO operational 183 IASI channel data, the experiment run used the newly selected 189 channels during the nighttime and 163 channels during the daytime after removing 26 IASI band 3 channels. Apart from the IASI channels, the same observational data were used for both the control and experiment runs.

In the first-guess departure analysis of the control and experiment runs, a significant difference in the humidity field was found when bias-corrected

sounders (i.e., ATOVS, AIRS, and CrIS) measurements and radiosonde observations were compared. The control run showed a small positive humidity bias, while the experiment run showed a negative bias whose absolute magnitude was significantly larger than the positive mean bias shown in the control run.

Results from the first-guess departure analysis of the humidity from radiosonde observations showed a control run with a large moisture bias, compared to the near-zero bias results from sounding sensor measurements. However, the moisture bias is substantially reduced in the experiment run, indicating the positive impact of the new IASI channels on the humidity field. The differences in the first-guess departure of the humidity field between the sounding sensors and radiosonde observations are likely due to the bias correction applied to the radiances measured by the sounding sensors. The global forecasts from the control and experiment runs also demonstrated that the humidity biases in the upper troposphere, shown in the control run, are substantially reduced by the use of the new IASI channels. The improvement in the humidity field is likely due to the use of different H_2O channels, which may better represent the upper-tropospheric humidity.

In addition to the global forecast impact, other trial experiment was conducted for a summer of 2015 (June 15–July 31, 2015), in order to evaluate the impact of the new IASI channels on the forecast accuracy of heavy precipitation

over East Asia. Because the heavy rainfall over East Asia tends to be induced by two different cloud systems, we evaluated the forecast accuracy of precipitation depending on the rain type (i.e., warm-type and cold-type heavy rainfall) derived by each cloud system as well as overall forecast accuracy of precipitation. Using the surface gauge data and the cloud-to-ground lightning data, the heavy rainfall events of 2015 were classified into two rainfall types, based on the characteristics of cloud-rain system of each rain type. Compared with the mean forecast scores (i.e., the TS and "Bias") of heavy precipitation at the control, the forecast accuracy of precipitation from the experiment run is improved by the use of new IASI channels. The forecast improvement is also shown in the warm-type heavy rainfall cases. In contrast, in case of cold-type heavy rainfall cases, the new IASI channels have an overall neutral impact on the forecast improvement of precipitation, compared to the control run.

One of caveats in this study stems from the use of diagonal component alone from the error covariance matrix, to give weights to water vapor profile and thus to alleviate the strong water vapor contributions to the channel selection. In addition, because IASI channels suitable for the NWP system considerably depend on various factors (e.g., sensor stability, performance of NWP model, and upgrade of radiative transfer model), the impact of ignoring the cross covariance should be assessed in the future.

References

Alduchov, O. A. and R. E. Eskridge, 1996: Improved Magnus form approximation of saturation vapour pressure. *J. Appl. Meteorol.*, **35**, 601–609, doi:http://dx.doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Andersson, E. and E. Hólm, P. Bauer, A. Beljaars, G. A. Kelly, A. P. McNally, A. J. Simmons, J.-N. Thépaut, and A. M. Tompkins, 2007: Analysis and forecast impact of the main humidity observing systems. *Q. J. R. Meteorol. Soc.*, **133**, 1473–1485.

Bates, J. J., and D. L. Jackson, 2001: Trends in upper-tropospheric humidity. *Geophys. Res. Lett.*, **28**, 1695–1698.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, doi:10.1038/nature14956.

Bormann, N., and P. Bauer, 2010: Estimates of spatial and interchannel observation error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Q. J. R. Meteorol. Soc.*, **136**, 1036–1050, doi:10.1002/qj.616.

Boukabara, S.-A., K. Garrett, and W. Chen, 2010: Global coverage of total precipitable water using a microwave variational algorithm. *IEEE Trans. Geosci. Remote Sens.*, **48**, 3608–3621, doi:10.1109/TGRS.2010.2048035.

Brown, A., S. Milton, M. Cullen, B. Golding, J. Mitchell, and A. Shelly, 2012: Unified modeling and prediction of weather and climate: A 25-year journey. *Bull. Amer. Meteor. Soc.*, **93**, 1865–1877, doi:http://dx.doi.org/10.1175/BAMS-D-12-00018.1.

Buck, A., 1981: New equation for computing vapor pressure and enhancement factor. *J. Appl. Meteorol.*, **20**, 1527–1532, doi: http://dx.doi.org/10.1175/1520-

0450(1981)020<1527:NEFCVP>2.0.CO;2.

Cameron, J., A. Collard, and S. English, 2005: Operational use of AIRS observations at the Met Office. Tech. Proceedings 14th Internat. TOVS Study Conf., Beijing, China, 25–31 May 2005.

Clark, H. L. and R. S. Harwood, 2003: Upper-tropospheric humidity from MLS and ECMWF reanalyses. *Mon. Wea. Rev.*, **131**, 542–555, doi:http://dx.doi.org/10.1175/1520-0493(2003)131<0542:UTHFMA>2.0.CO;2.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Q. J. R. Meteorol. Soc.*, **139**, 1445–1461, doi:10.1002/qj.2054.

Collard, A. D., 2007: Selection of IASI channels for use in numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **133**, 1977–1991, doi: 10.1002/qj.178.

Collard, A. D. and A. P. McNally, 2009: The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF. *Q. J. R. Meteorol. Soc.*, **135**, 1044–1058, doi: 10.1002/qj.410.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, **131**, 1759–1782.

Deblonde, G., J. F. Mahfouf, B. Bilodeau, and D. Anselmo, 2007: One-dimensional variational data assimilation of SSM/I observations in rainy atmospheres at MSC. *Mon. Wea. Rev.*, **135**, 152–172, doi: http://dx.doi.org/10.1175/MWR3265.1.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V.

Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597, doi: 10.1002/qj.828.

Dessler, A. E., and S. C. Sherwood, 2000: Simulations of tropical upper tropospheric humidity. *J. Geophys. Res.*, **105**, 20155–20163.

Dirksen, R. J., M. Sommer, F. J. Immler, D. F. Hurst, R. Kivi, and H. Vömel, 2014: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde. *Atmos. Meas. Tech.*, **7**, 4463–4490, doi:10.5194/amt-7-4463-2014.

Draper, D. W., D. A. Newell, F. J. Wentz, S. Krimchansky, and G. M. Skofronick-Jackson, 2015: The Global Precipitation Measurement (GPM) Microwave Imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, **8**, 3452–3462, doi:10.1109/JSTARS.2015.2403303.

English, S. J., 1999: Estimation of temperature and humidity profile information from microwave radiances over different surface types. *J. Appl. Meteorol.*, **38**, 1526–1541, doi: http://dx.doi.org/10.1175/1520-0450(1999)038<1526:EOTAHP>2.0.CO;2.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J. Rayer, C. Poulsen, F. W. Saunders, and J. R. Eyre, 2000: A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Q. J. R. Meteorol. Soc.*, **126**, 2911–2931, doi: 10.1002/qj.49712656915.

EUMETSAT, 2012: IASI Level 1 Product Guide, EUM/OPS-EPS/MAN/04/0032, [Available online at http://oiswww.eumetsat.org/WEBOPS/eps-pg/IASI-L1/IASIL1-PG-0TOC.html].

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys.*

Res., **99**, 10143–10162, doi:10.1029/94JC00572.

Gambacorta, A., and C. Barnet, 2013: Methodology and information content of the NOAA NESDIS operational channel selection for the Cross-Track Infrared Sounder (CrIS). *IEEE Remote Sensing.*, **51**, 3207–3216, doi: 10.1109/TGRS.2012.2220369.

Goff, J. A., 1957: Saturation pressure of water on the new Kelvin temperature scale. *Trans. Am. Soc. Heat. Vent. Eng.*, 347–354.

Goff, J. A., and S. Gratch, 1946: Low-pressure properties of water from -160 to 212 F. *Trans. Am. Soc. Heat. Vent. Eng.*, **52**, 95–121.

Harris, B. A., and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q. J. R. Meteorol. Soc.*, **127**, 1453–1468, doi: 10.1002/qj.49712757418.

Hilton, F., A. C. Atkinson, S. J. English, and J. R. Eyre, 2009: Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. *Q. J. R. Meteorol. Soc.*, **135**, 495–505, doi: 10.1002/qj.379.

Hilton, F. I., S. M. Newman, and A. D. Collard, 2012: Identification of NWP humidity biases using high-peaking water vapour channels from IASI. *Atmos. Sci. Let.*, **13**, 73–78, doi:10.1002/asl.366.

Hong, S.-Y., 2004: Comparison of heavy rainfall mechanisms in Korea and the Central US. *J. Meteor. Soc. Japan*, **82**, 1469-1479.

Huffman, G., D. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. Nelkin, and P. Xie, 2013: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Document, version 4.5. NASA (30 pp.).

Hyland, R. W., and A. Wexler, 1983: Formulations for the thermodynamic properties of the saturated phases of H2O from 173.15 K to 473.15 K. *ASHRAE Trans.*, **89**, 500–519.

Immler, F. J., and M. Sommer, 2011: Brief description of the RS92 GRUAN data product (RS92-GDP). GRUAN Tech. Document GRUAN-TD-4, 17 pp.

Joo, S., J. R. Eyre, and R. Marriott, 2013: The impact of Metop and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Mon. Wea. Rev.*, **141**, 3331–3342, doi: 10.1175/MWR-D-12-00232.1.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 342 pp.

Kelly, G., and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the Global Observation System through Observing System Experiments. *ECMWF Newsletter*, **113**, 16–28.

Korea Meteorological Administration, 2012: Joint WMO technical progress report on the global data processing and forecasting system and numerical weather prediction research activities for 2012, 20 pp.

Kwon, E.-H., J. Li, J. Li, B.J. Sohn, E. Weisz, 2012: Use of total precipitable water classification of a priori error and quality control in atmospheric temperature and water vapor sounding retrieval. *Adv. Atmos. Sci.*, **29**, 263–273, doi: 10.1007/s00376-011-1119-z.

Kwon, E.-H., B.J. Sohn, W.L. Smith, J. Li, 2012: Validating IASI temperature and moisture sounding retrievals over East Asia using radiosonde observations. *J. Atmos. Ocean. Technol.*, **29**, 1250–1262, doi: 10.1175/JTECH-D-11-00078.1.

Lee, J.-W., S.-O. Han, and K.-Y. Chung, 2012: The Improvement of forecast accuracy of the unified model at KMA by using an optimized set of physical options (in Korean with English abstract). *J. Korean Meteor. Soc.*, **22**, 345–356.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177–1194, doi:10.1002/qj.49711247414.

Lorenc, A. C., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders, 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126**, 2991–3012, doi:10.1002/qj.49712657002.

Lorenc, A. C., and R. T. Marriott, 2014: Forecast sensitivity to observations in the Met Office global numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **140**, 209–224, doi: 10.1002/qj.2122.

Martinet, P., L. Lavanant, N. Fourrié, F. Rabier, and A. Gambacorta, 2014: Evaluation of a revised IASI channel selection for cloudy retrievals with a focus on the Mediterranean basin. *Q. J. R. Meteorol. Soc.*, **140**, 1563–1577, doi:10.1002/qj.2239.

Matricardi, M., F. Chevallier, G. Kelly, and J.-N. Thépaut, 2004: An improved general fast radiative transfer model for the assimilation of radiance observations. *Q. J. R. Meteorol. Soc.*, **130**, 153–173, doi: 10.1256/qj.02.181.

Met Office, 1998: Unified Model User Guide, p200.

Met Office, 2011: Joint WMO technical progress report on the global data processing and forecasting system and numerical weather prediction research activities for 2011. WMO GDPFS/NWP Rep., Available online: http://www.wmo.int/pages/prog/www/Documents/PublicWeb/www/gdpfs/GDPFS-NWP_Annualreports11/2011_UK.pdf.

Migliorini, S., 2015: Optimal ensemble-based selection of channels from advanced sounders in the presence of cloud. *Mon. Wea. Rev.*, **143**, 3754–3773, doi: 10.1175/MWR-D-14-00249.1.

Miloshevich, L. M., H. Vömel, D. Whiteman, B. Lesht, F. J. Schmidlin, and F. Russo, 2006: Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation. *J. Geophys. Res.*, **111**, doi:10.1029/2005JD006083.

Moradi, I., B. Soden, R. Ferraro, P. Arkin, and H. Vömel, 2013: Assessing the quality of humidity measurements from global operational radiosonde sensors. *J. Geophys. Res. Atmos.*, **118**, 8040–8053, doi:10.1002/jgrd.50589.

Murphy, D. M., and T. Koop, 2005: Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Q. J. R. Meteorol. Soc.*, **131**, 1539–1565, doi: 10.1256/qj.04.94.

Nash, J., T. Oakley, H. Vömel, and W. Li, 2011: WMO Intercomparisons of high quality radiosonde systems. Tech. Rep., World Meteorological Organization, Yangjiang, China.

Rabier, F., N. Fourrié, D. Chafaï, and P. Prunet, 2002: Channel selection methods for Infrared Atmospheric Sounding Interferometer radiances. *Q. J. R. Meteorol. Soc.*, **128**, 1011–1027, doi: 10.1256/0035900021643638.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global 4-Dimensional data assimilation system. *Q. J. R. Meteorol. Soc.*, **133**, 347–362, doi:10.1002/qj.32.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Centre's spectral statistical interpolation analysis scheme. *Mon. Wea. Rev.*, **120**, 1747–1763, doi: http://dx.doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Pavelin, E. G., S. J. English, and J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **134**, 737–749, doi: 10.1002/qj.243.

Petersen, W. A., and S. A. Rutledge, 1998: On the relationship between cloud-to-ground lightning and convective rainfall. *J. Geophys. Res.*, **103**, 14 025–14 040.

Rodgers, C. D., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Ser. Atmos. Oceanic Planet. Phys., Vol. 2, World Scientific,

256pp.

Saunders, R. W., T. A. Blackmore, B. Candy, P. N. Francis, and T. J. Hewison, 2013: Monitoring satellite radiance biases using NWP models. *IEEE Trans. Geosci. Remote Sens.*, **51**, 1124–1138, doi: 10.1109/TGRS.2012.2229283.

Seity, Y., S. Soula, and H. Sauvageot, 2001: Lightning and precipitation relationship in coastal thunderstorms. *J. Geophys. Res.*, **106**, 22 801–22 816

Siméoni, D., C. Singer, and G. Chalon, 1997: Infrared atmospheric sounding interferometer. *Acta Astron.*, **40**, 113–118, doi: 10.1016/S0094-5765(97)00098-2.

Simmons, A. J., A. Untch, C. Jakob, P. Kallberg, and P. Unden, 1999: Stratospheric water vapour and tropical tropopause temperatures in ECMWF analyses and multi-year simulations. *Q. J. R. Meteorol. Soc.*, **125**, 353–386, doi:10.1002/qj.49712555318.

Smith, F. I., 2014: Improving the information content of IASI assimilation for numerical weather prediction. PhD Thesis, Department of Physics and Astronomy, University of Leicester.

Sohn, B. J., G. H. Ryu, H. J. Song, and M. L. Ou, 2013: Characteristic Features of Warm-Type Rain Producing Heavy Rainfall over the Korea Peninsula Inferred from TRMM Measurements. *Mon. Wea. Rev.*, **141**, 3873-3888.

Song, H.-J., and B. J. Sohn, 2015: Two heavy rainfall types over the Korean peninsula in the humid East Asian summer environment: A satellite observation study. *Mon. Wea. Rev.*, **143**, 363-382, doi: 10.1175/MWR-D-14-00184.1.

Soriano, L. R., F. Pablo, and E. G. Diez, 2001: Relationship of precipitation and cloud-to-ground lightning in the Iberian Peninsula. *Mon. Wea. Rev.*, **129**, 2998 – 3003.

Sun, B., A. Reale, D. J. Seidel, and D. C. Hunt, 2010: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics. *J. Geophys. Res.*, **115**, D23104, doi:10.1029/2010JD014457.

Susskind, J., J. Blaisdell, L. Iredell, and F. Keita, 2011: Improved temperature sounding and quality control methodology using AIRS/AMSU data: The AIRS Science Team version 5 retrieval algorithm. *IEEE Trans. Geosci. Remote Sens.*, **49**, 883–907.

Wang, J., L. Zhang, A. Dai, F. Immler, M. Sommer, and H. Vömel, 2013: Radiation dry bias correction of Vaisala RS92 humidity data and its impact on historical radiosonde data. *J. Atmos. Oceanic Technol.*, **30**, 197–214, doi: http://dx.doi.org/10.1175/JTECH-D-12-00113.1.

Warner, T. T., 2011: *Numerical Weather and Climate Prediction*. Cambridge University Press, 548 pp.

Weston, P., 2011: Progress towards the Implementation of Correlated Observation Errors in 4D-Var. Forecasting Research Technical Report 560, Met Office, Exeter, UK.

Weston, P. P., W. Bell, and J. R. Eyre, 2014: Accounting for correlated error in the assimilation of high-resolution sounder data. *Q. J. R. Meteorol. Soc.*, **140**, 2420–2429, doi: 10.1002/qj.2306.

Weston, P., E. G. Pavelin, and A. D. Collard, 2013: NWP SAF Met Office 1D-Var User Manual. NWPSAF-MO-UD-006.

국문초록

자료동화성능 개선을 하기에 앞서 현재 한국 기상청 및 영국 기상청에서 수치예보모델로 사용하고 있는 통합수치모델 (Unified Model) 내 자료동화과정을 통해 생산되는 초기장 (analysis)의 성능을 진단하였다. 뿐만 아니라 상호 비교를 위하여 유럽중기예보센터 (ECMWF)에서 제공하는 재분석장 (ERA-Interim)도 동일한 방법으로 진단하였다. 진단에 사용한 관측자료는 독일기상청에서 제공하는 GRUAN 라디오존데 자료이며, 2012년 1월부터 12월에 해당하는 자료를 확보하여 사용하였다.

온도의 경우 수치모델 종류에 상관없이 연직으로 평균오차 ±0.5 K 미만, 평균제곱근 오차는 1.0 K 미만으로 비교적 정확한 성능을 보이는 반면에 습도의 경우 상층에서 모델 종류에 따라 오차가 크게 변동하였다. 특히 유럽중기예보센터의 경우 상층에서 최대 6% 습윤오차가 있었으며, 평균제곱근 오차에서는 800 hPa 근처에서 15%에 이르렀다. 동일한 수치모델을 사용하고 있음에도 불구하고 한국기상청 및 영국기상청의 경우 평균 제곱근 오차는 700 hPa 부근에서 각각 14%, 12%에 이르렀다. 수증기량에 따른 분석에서는 습도의 경우 수증기량이 많은 습한 환경에서는 수치모델 종류에 상관없이 상층에서 습윤 오차가 존재하였다. 그리고 평균제곱근 오차는 라디오존데 관측자료로부터 생산된 수증기량에 따른 표준오차에 비례하는 양상을 보였으며, 이는 수증기 변동성이 큰 환경을 모의하는데 있어서 수치예보모델의 한계점이 있

음을 의미한다.

이러한 통합수치모델에서 생산되는 분석장 오차를 개선하기 위하여 IASI 채널 선택 연구를 진행하였다. 현재 통합수치모델 자료동화과정에서는 과거에 선택된 IASI 183개 채널 관측자료가 사용되고 있으나, 이후로 수치모델성능 및 자료동화기술이 향상되었기 때문에 현재 시스템에 적합한 IASI 채널을 다시 선택하였다. 기존에 Rodgers (2000)에서 제안한 선택방법이 가지는 문제점을 해결하고자 일차원 변분법을 이용한 Channel Score Index (CSI) 방법을 이용하여 최종적으로 200개 채널을 선택하였다. 기존에 사용되고 있던 IASI 183개 채널과 비교해본 결과 새로 선정된 200채널 내에 오존 및 다수의 하층 및 상층 수증기 채널이 추가되었으며, 단파적외영역 내 다수 채널이 포함되었다.

새로 선정된 IASI채널이 전지구 예보정확도에 미치는 영향을 알아보기 위하여 영국 기상청 자료동화시스템에 적용한 결과, 전반적인 수치예보정확도에 있어서 개선은 없었다. 하지만 기존 수치모델 예보장에 존재하던 상층 습윤오차가 크게 감소하였으며, 72시간 예보장까지 유지되는 모습을 보였다. 이러한 상층 습윤 오차감소는 새로 선정된 IASI 채널 내 추가된 상층 수증기 채널의 효과라고 할 수 있다.

전지구 예보 정확도뿐만 아니라 새로 선정된 IASI 채널의 동아시아 지역 여름철 강수예보정확도에 미치는 영향에 대해서도 알아보기 위하여 한국 기상청 자료동화시스템을 이용하여 실험을 수행하였다. 동아시아 여름철에 내리는 강한 강수의 경우 다른 지역에서 내리는 강수와 다른 특징을 가지고 있

다. 일반적으로 불안정한 대기 상태에서 급격하게 높게 발단한 구름에 의해 강한 강수가 발생하는데, 동아시아 지역의 경우 비교적 낮은 구름에서도 강한 강수가 발생하는 경우가 빈번히 발생한다. 이러한 동아시아 지역에서 특징적으로 발생하는 강수 예보 정확도에 미치는 영향을 알아보기 위해 강수유형 (i.e., warm-type과 cold-type)을 분류하여 분석하였다. 새로 선정된 IASI 채널을 적용했을 경우 전반적인 동아시아 강수 예보 정확도는 개선되었으며, 특히 과대모의 되고 있던 강수영역이 크게 개선되었다. 강수유형에 따른 결과에서도 동아시아 지역에서 특징적으로 발생하는 warm-type 강수에서 예보 정확도 개선을 볼 수 있었다.

주요어: 적외 초분광 센서, 채널선택, 자료동화, 강수, 동아시아

학 번: 2009-20367